# Dynamic Modelling of DNA Repair Pathway at the Molecular Level: A New Perspective

**Lecca, P. & Ihekwaba-Ndibe, A. E. C.**

# Dynamic Modelling of DNA Repair Pathway at the Molecular Level: A New Perspective

Paola Lecca[1]* and Adaoha E. C. Ihekwaba-Ndibe[2]*

[1]Faculty of Computer Science, Free University of Bozen-Bolzano, Bolzano, Italy, [2]Faculty of Health and Life Sciences, Coventry University, Coventry, United Kingdom

DNA is the genetic repository for all living organisms, and it is subject to constant changes caused by chemical and physical factors. Any change, if not repaired, erodes the genetic information and causes mutations and diseases. To ensure overall survival, robust DNA repair mechanisms and damage-bypass mechanisms have evolved to ensure that the DNA is constantly protected against potentially deleterious damage while maintaining its integrity. Not surprisingly, defects in DNA repair genes affect metabolic processes, and this can be seen in some types of cancer, where DNA repair pathways are disrupted and deregulated, resulting in genome instability. Mathematically modelling the complex network of genes and processes that make up the DNA repair network will not only provide insight into how cells recognise and react to mutations, but it may also reveal whether or not genes involved in the repair process can be controlled. Due to the complexity of this network and the need for a mathematical model and software platform to simulate different investigation scenarios, there must be an automatic way to convert this network into a mathematical model. In this paper, we present a topological analysis of one of the networks in DNA repair, specifically homologous recombination repair (HR). We propose a method for the automatic construction of a system of rate equations to describe network dynamics and present results of a numerical simulation of the model and model sensitivity analysis to the parameters. In the past, dynamic modelling and sensitivity analysis have been used to study the evolution of tumours in response to drugs in cancer medicine. However, automatic generation of a mathematical model and the study of its sensitivity to parameter have not been applied to research on the DNA repair network so far. Therefore, we present this application as an approach for medical research against cancer, since it could give insight into a possible approach with which central nodes of the networks and repair genes could be identified and controlled with the ultimate goal of aiding cancer therapy to fight the onset of cancer and its progression.

Keywords: DNA damage, DNA repair genes, dynamical networks, ODE models, parametric sensitivity analysis, centrality measure analysis

# 1 INTRODUCTION

DNA molecules packaged in our chromosomes carry our genetic blueprint, and their preservation is essential for the coordination of cellular function and organization of life (Branzei and Foiani, 2008). As DNA is the repository of our genetic information, we would expect its structure to be highly stable. This is not the case for the DNA (Reed and Waters, 2005). Damage to DNA is a constant threat (Lindahl, 1993; Alberts, 2015). The DNA molecule is intrinsically reactive, as it is very susceptible to chemical and physical factors, which can lead to DNA lesions, such as base loss, base modification, and double-strand DNA breaks (Hoeijmakers, 2009; Çağlayan and Wilson, 2015; Ross and Truant, 2016; Yadav et al., 2020). Physiological conditions such as oxygen-rich, aqueous, or pH 7.4 (Lindahl, 1993) as well as chemical events such as hydrolysis and exposure to reactive oxygen species (ROS) or other reactive metabolites can damage DNA. Exogenous chemicals or endogenous metabolic processes trigger chemical reactions. Although exogenous stressors can be extremely powerful, endogenous threats are constant and unabating. It is estimated that a single cell experiences up to $10^5$ spontaneous or induced DNA lesions per day (Lindahl, 1993; Bont, 2004; Kovalchuk, 2016; Chatterjee and Walker, 2017).

DNA damage has far-reaching consequences, such as preventing RNA polymerase from transcribing the correct messenger RNA sequence to produce the correct protein. In the longer term, cellular malfunctions such as cancer initiation, inborn defects, and ageing that result after damaged DNA replicates are examples of unpredictable long-term consequences of DNA damage; as base misincorporation causes mutations which alter the genetic code (Reed and Waters, 2005). Therefore, a coordinated response to DNA damage is necessary in order to ensure cellular viability and prevent diseases. Cells, fortunately, possess a robust system of mechanisms that function together to reduce the adverse consequences of DNA damage and ensure that their genetic information is faithfully replicated, thus maintaining the integrity of their genome (Ganai and Johansson, 2016). This coordinated effort, known as DNA damage response (DDR) operates by sensing and signalling the genotoxic events, and the damage is then resolved either by DNA repair machineries, or cell death if DNA cannot be repaired. DNA repair functions as part of the DNA damage response (DDR) (Liu, 2001; Chatterjee and Walker, 2017; Reed and Waters, 2005; Hoeijmakers, 2001, 2009).

DNA repair has so far been shown to exist in both prokaryotic and eukaryotic organisms, with over 150 proteins directly involved in safeguarding the genome (Sancar et al., 2004; Friedberg et al., 2005; Wood et al., 2005; Yousefzadeh et al., 2021). DNA repair processes restore DNA back to its normal sequence and structure after damage (Friedberg et al., 2006), and are characterised traditionally by the type of damage they repair. There are five major DNA repair pathways available to cells to deal with DNA damage burdens. Each of these processes recognises a particular type of DNA lesions, and together work in preventing mutagenesis. They include 1) direct reversal repairs that repairs lesion induced mainly by alkylating agents, 2) Base excision repair (BER), for small base modifications like single-strand breaks (SSBs) and non-bulky damaged DNA bases, 3) Nucleotide excision repair (NER), that corrects bulky, helix-distorting DNA lesions, 4) mismatch repair (MMR), that repairs base-base mismatch and insertion or deletion loops (IDLs), 5) Recombinational repair, which is divided into non-homologous end joining (NHEJ) and homologous recombination repair (HR), both of which repairs DNA double-strand breaks (DSBs). Other types of DSB repairs include alternative non-homologous end-joining (alt-NHEJ, MMEJ) and translesion synthesis (TLS), which operates as a tolerance mechanism for DNA damage (Jackson and Bartek, 2009; Hosoya and Miyagawa, 2014; Li et al., 2021).

For frequently occurring DNA damage, direct reversal of DNA damage by specialised proteins is the most efficient and most straightforward method of DNA repair. However, this approach is only used by a small proportion of DNA repair types. Most damage to DNA is repaired by the removal of damaged bases and is followed by resynthesis of the removed/excised region (replacement) (Cooper, 2000). The pathways involved in the removal of base damage are base excision repair (BER), nucleotide excision repair (NER) and mismatch repair (MMR) (Cooper, 2000). The rest of the pathways repair damage to DNA structure/backbone. DNA damage can cause breaks in the DNA backbone, single-strand breaks in one strand, or double-strand breaks on both strands. Single-strand breaks are repaired by mechanisms sharing common steps in the BER pathway; however, DSBs are especially harmful as, by definition, no unbroken complementary strand exit which can serve as a template for repair when both strands break (Bennett et al., 1993; Friedberg et al., 2006). For cells with DNA already replicated prior to cell division, the duplicate copy can easily supply the missing information. So in these cells, DSBs can be repaired by HR, involving the exchange of DNA strands.

Even so, very efficient repair mechanisms can sometimes fail to provide a clean template for DNA synthesis. Replication errors can make it past these mechanisms, as DNA repair can also undergo mutations and become dysregulated. DNA repair gene mutations have been known to cause a variety of rare inherited human syndromes. Some of which include premature ageing phenotypes, increased sensitivity to ionising radiation exposure, and increased cancer risk (Friedberg et al., 2006; Lok and Powell, 2012; Carusillo and Mussolino, 2020). It has also been found that inherited defects in each of the DNA repair pathways are associated with distinct genome instability syndromes (Yousefzadeh et al., 2021), syndromes characterised by developmental defects (Bouwman and Jonkers, 2012; Ghosal and Chen, 2013; Wolters and Schumacher, 2013; Wood, 2018).

The dysregulation of DNA repair gene networks underlies many human genetic diseases that affect a wide range of body systems but all share a common trait, predisposition to cancer (Chatterjee and Walker, 2017). Almost all human cancers are spontaneous, not inherited, and are caused by environmental or genetic factors. It is of great public health interest to determine which genetic variations increase cancer risk in normal populations, and DNA repair genes are likely contenders.

Therefore, elucidating the molecular mechanism behind DNA repair defects may provide a framework for understanding the complex pattern of genetic variations that contribute to spontaneous human cancers.

In this study, we demonstrate how to translate a network (mathematically definable as a hyper-graph) into a set of first-order differential equations of the mass action law type. Once the model has been established, we present its numerical solution and carry out a sensitivity analysis of its kinetic rates, whose numerical values are mostly unknown. The results of the analysis of network dynamics complement those produced by the calculation of centrality measures, and together they produce a set of genes of similar biological interest, and in perspective also of medical interest, due to their characteristics of topological centrality and vulnerability to stimuli. The paper is organized as follows: in **Section 2** we describe the mechanisms of double-strand break repair pathway homologous recombination repair necessary to understand and interpret the results of the computational analysis, in **Section 3** we describe the rules on which the automatic translation of the network into a rate equation system is based and the methods of sensitivity analysis of the model. In **Section 4** we present the results of the analysis, and, finally, in **Section 5**, we draw some conclusions.

# 2 DOUBLE-STRAND BREAK REPAIR PATHWAY HOMOLOGOUS RECOMBINATION REPAIR

DSBs are the most serious DNA damage, as both DNA strands are impaired simultaneously. Therefore, due to the magnitude of differing factors leading to DSBs, the effectiveness of their repair is crucial for cell survival and the functioning and prevention of DNA fragmentation, chromosomal translocation and deletion. DSBs can be repaired in mammalian cells by NHEJ, HR, and single-strand annealing (SSA). Unrepaired SSBs result in much more cytotoxic DSBs formation during the S-phase progression of the cell cycle (Kennedy and D'Andrea, 2006). Homologous recombination is a process by which DSBs are repaired through the alignment of homologous sequences of DNA (Dietlein and Reinhardt, 2014) and occurs primarily during the late S to G2 phase of the cell cycle (Cerbinskaite et al., 2012; Chatterjee and Walker, 2017).

Homologous recombination is the second major DSB repair pathway and requires a second, homologous DNA sequence to function as donor template. There are two phases to this process, the first phase triggered by sensor proteins that belong to the MRN complex, and the second phase by the stimulation of resection steps, initiated in the first phase and subsequently extended. HR generally involves the following stages:

1. DSBs are recognised and sensed by the MRN complex (Kim et al., 1994), which activates ATM kinase, initiating the DSB end resection steps, where CtBP-interacting protein (CtIP) and the MRN complex work together to generate single-strand DNA (ssDNA) at the DSB ends ((Zhao et al., 2020).

2. The exposed ssDNA is recognised by and coated with DNA replication protein A (RPA) complex, which recruits the major homologous recombination regulator RAD52 to the site to facilitate HR repair (Maréchal and Zou, 2014; Rossi et al., 2021).

3. The nucleoprotein filament RAD51, is then assembled, mediated by BReast CAncer type 2 susceptibility protein (BRCA2), to replace RPA on ssDNA to perform homology sequence searching and strand invasion (Kowalczykowski, 2015).

4. DSBs are then restored by branch migration, DNA synthesis, ligation, and resolution of Holliday junctions (Zhao et al., 2020).

Following the recognition and sensing of DSBs, a process known as DNA end resection is activated, a critical function in HR (Liu and Huang, 2016; Zhao et al., 2020). DNA end resection catalyses the nucleolytic degradation of the broken ends of DSBs (by the CtIPMRN complex) in the 5′ to 3′ direction generating 3′ single-stranded DNA (ssDNA). The 3′ ssDNA then provides a platform for the recruitment of HR repair-related proteins (Huertas, 2010; Liu and Huang, 2016; Zhao et al., 2020). Following the generation of ssDNA, downstream nucleases and helicases, such as exonuclease 1 (EXO1) or DNA replication ATP-dependent helicase/nuclease DNA replication helicasenuclease 2 (DNA2) and Bloom syndrome protein (BLM), are conscripted to extend the 3' ssDNA for HR repair (Huertas and Jackson, 2009; Yun and Hiom, 2009; Zhao et al., 2020). The identities of these DNA helicases and nucleases are yet to be clearly defined in humans (as in yeast), partly because there are many candidate proteins. Although five RecQ helicase homologs have so far been identified in yeast (Bloom helicase [BLM], Werner helicase/nuclease [WRN], RECQ1, RECQ4, and RECQ5) (Chu and Hickson, 2009; Lu and Davis, 2021), convincing evidence point up BLM in resection (Gravel et al., 2008; Nimonkar et al., 2008). Following resection, the exposed single-strand DNA (ssDNA) is recognised and bound by RPA complex for protection.

RPA plays a significant role in coordinating DNA resection processes and simultaneously preserving the integrity of the resultant ssDNA (Sun et al., 2019). RPA is a heterotrimeric ssDNA binding protein essential to nearly all DNA processing events and associates with ssDNA with very high affinity (Kd sim 109–1010 M) (Maréchal and Zou, 2014). It is comprised of three protein subunits, RPA70, RPA32 and RPA14 and contains multiple oligonucleotideoligosaccharide (OB)-folds that interact with both ssDNA and proteins (Kim et al., 1994; Fanning, 2006; Feldkamp et al., 2014; Maréchal and Zou, 2014). RPA is flexible (Brosey et al., 2013). Its versatile nature allows it to coordinate the recruitment, activation and exchange of many proteins whose combined activities allow for the protection and propagation of eukaryotic genomes (Maréchal and Zou, 2014). How multiple RPAs associate on ssDNA and coordinate its vast array of processes remains to be determined (Sun et al., 2019). However, a critical feature of RPA is that, though it can bind nucleic acids with very high affinity, it can

**TABLE 1 |** Interactive graphical user interface of NADS software showing the options and the task concerning the generation of ODE equations and their solution.

```
#######################################################
Welcome to Network Analyzer and Dynamics Simulator
#######################################################

Which network do you want to analyze?

1: Homologous recombination interactions
2: Test example
3: Exit

Selection: 1

##################################
Principal network characteristics
##################################

Number of nodes: 25
Number of edges: 250

Types of interactions and their frequency:

        controls expression of 1
        controls state change of 134
        in catalysis with 41
        interacts with 74

###############################
Centrality measures calculation
###############################

Degree

Betweennes

Clustering coefficient

Eingenvector centrality

Vibrational centrality
... pseudo-inverse of graph Laplacian ...

Subgraph centrality

Information centrality
... pseudo-inverse for information centrality estimation ...
Export results to files ...

Select last decile nodes from the centralities measure distributions ...
Nodes highly scoring by Clustering coefficient : empty set.

Select first decile nodes from the centralities measure distributions ...

... save the results of selection in barplots.

Saving plots of centralities distributions.
```

**TABLE 2 |** Interactive graphical user interface of NADS software showing the options and the task concerning the analysis of network topology.

```
##################################################
                Build and solve network dynamics
##################################################

Network size reduction: the union of the nodes of 99-th percentile
of centrality distributions will be selected.

Do you want to reduce the size of the network to analyze?
(It is highly recommended to select Yes if you want to reduce
the computational times.)

1: Yes
2: No
3: Exit

Selection: 2

The number of nodes of the network is:  25
The number of edges of the network is:  250

Start to convert this network into a set of ordinary differential
equations (ODE) ...

The initial values of the variables are selected randomly in the
range [1, 100].
Do you want to change this range?

1: Yes
2: No
3: Exit

Selection: 2

The input file does not contain information on the kinetic rate constants:
random values from a given range will be selected (default is [0,0.01]).
Do you want to change this range?

1: Yes
2: No
3: Exit

Selection:  2

... script dynamics.R has been created.......: 11.801 sec elapsed
Solve the ODE of the network dynamics ...: 11.961 sec elapsed
```

easily be displaced by other enzymes for further downstream processing (Sun et al., 2019).

When ssDNA length is sufficient for HR repair, the end resection process is terminated (Zhao et al., 2015). Although the regulation of DNA end resection termination are not yet clearly understood, several studies suggest that under physiological conditions, end resection is terminated by RAD51-RPA switching (Zhao et al., 2015). This switching is regulated by BRCA2-DSS1. DSS1 - SEM1 in yeast - is a small,

highly acidic protein that competes with ssDNA, by mimicking ssDNA in order to remove RPA from the genuine ssDNA (Zhao et al., 2015; Stefanovie et al., 2019; Le et al., 2020; Rossi et al., 2021). The DSS1 then binds to BRCA2 in order to facilitate RAD51 filament formation (Liu et al., 2010; Stefanovie et al., 2019; Rossi et al., 2021). DSS1 does not seem to bind DNA on its own but appears to enhance ssDNA binding activities of BRCA2 and RAD52 to promote DSB repair (Zhao et al., 2015). BRCA2 then recruits RAD51 to complete the switch

(Zhao et al., 2015; Rossi et al., 2021). For cells with DNA already replicated prior to cell division, RAD51 will oligomerise and form a nucleoprotein filament on the resected, single-stranded DNA (ssDNA) end of the DSB, and search for the homologous DNA sequence on the undamaged sister chromatid, performs strand exchange (invasion), and produce a joint molecule called a D-loop (Rossi et al., 2021). DNA polymerase will then use the homologous DNA strand as a template from the D-loop, and the 3′-end of the broken DNA strand as a primer to commence DNA repair synthesis (Rossi et al., 2021). The other end of the double-strand break is then apprehended by RAD52, joining it to the D-loop, through the annealing process, causing the displaced strand to act as a template for the second strand synthesis (Rossi et al., 2021). When DNA synthesis is complete, the D-loops are then dissociated by RAD54, a protein that interacts with RAD51 to promote branch migration, or interacts with helicases like BLM (van Brabant et al., 2000; Bugreev et al., 2006; Kawale and Sung, 2020; Rossi et al., 2021). DNA is further extended by DNA polymerase, annealed to the ssDNA part of the second broken DNA, gap filled and finally restored (Rossi et al., 2021). In **Figure 1** we summarise what is described in this section about the DSB signalling mechanisms.

Homologous recombination is able to repair DSBs error-free using the undamaged sister chromatid (Dietlein and Reinhardt, 2014). As the accuracy of homologous recombination repair is important for DSBs (Sugiyama and Kantake, 2009), if it is impaired, chemotherapeutic opportunities may arise (Huang and Zhou, 2021).

# 3 GRAPH REPRESENTATION AND MATHEMATICAL MODEL

We considered HDR through Homologous Recombination (HRR) network as available in Pathways Commons (Cerami et al., 2010) in the SIF (Simple Interaction Format) format at the link in the reference (Orlic-Milacic, 2015). See these data also reported in **Supplementary Tables S1–S3**.

We implemented an R script, that takes as input the HR network and is able to.

- analyse the topology of the network through the calculation of standard and new centrality measures. The standard node centrality measures considered in this study are the degree (in-, out- and total), the betweenness, the clustering coefficient, the eingenvector centrality, the vibrational centrality, the subgraph centrality, and the information centrality (see (Marsden, 2005; Koschützki and Schreiber, 2008; Ghasemi et al., 2014; Wang et al., 2014; Fornito et al., 2016; Jalili et al., 2016; Ashtiani et al., 2018) for a concise but comprehensive report on the meaning and the use of these measure in molecular biology). We considered also a new centrality measures, such as vibrational centrality, introduced by Estrada in (Estrada and Hatano, 2010) that we will discuss in more detail in the next section (we also

refer the reader to (Lecca and Re, 2019) for a review on vibrational centrality); for the reader's convenience, we list the definition of these centrality indices in **Supplementary Table S4**, that are also extensively covered in many textbooks on graph theory, and in various articles in the applied sciences. We refer the reader to Estrada's numerous works, a comprehensive compendium of which can be found in the book (Estrada and Hatano, 2010; Estrada, 2011);

- automatically generate a system of rate equations, specifically first order mass action differential equations, describing the dynamics of the network.

and a R script implementing parametric sensitivity analysis of the dynamics model.

In the dynamics model, by default the kinetic rate constants $k$ a well as the initial values of the proteins and molecules concentrations are set equal to random values in fixed ranges. Nevertheless these ranges can be modified by the user as shown by the interactive console output reported in **Table 1**. However, we note that the interval of definition of the uniform distribution cannot exceed the maximal range of parameter variability within which the system of rate equations has a solution. We refer the reader to a previous work of us (Lecca et al., 2016) for more details on this.

In a first experiment, the initial values of the proteins concentration (not experimentally known) has been drawn randomly in a range [1, 100] a. u., and the simulation time interval was [0, 10] a. u. In a second experiment, the numerical simulation of the model was performed by assigning an initial quantity between 18 and 20 (expressed in arbitrary units a. u.) to each node and for $t \in [0, 1400]$ (in arbitrary units). The solution of the system of 25 differential equations converges for values of rate constants in the range []0, 0.01 a. u. in the first experiment, and in the range []$10^{-6}$, $10^{-5}$ a.u. in the second experiment. To the best of our knowledge, the *in vivo* concentrations of the proteins that are part of the HR network are not known. Combined with the lack of knowledge of the values of rate constants, this means that we cannot assign units to these values that reflect *in vivo* kinetics. However, the initial values of the protein concentrations have been chosen from the order of magnitude at which various *in vitro* experiments operate and express these concentrations in nM and time in seconds. The literature we have referred to includes the works of (Yang et al., 2010; Nguyen et al., 2013; Foertsch et al., 2019).

We changed the parameters one at a time while keeping the values of the others fixed. Since for each parameter $p_h$ ($h$ ranges from 1 to the number of parameters in the equations), we sampled $N_P$ values, and consequently we performed $N_P$ model simulations. Let us denote with $x_s(t)$, ($s = 1, 2, \ldots, d$) the time series expressing the numerical solutions of the rate equations, where $d$ is the number of the proteins in the network. The index of sensitivity of $x_s(t)$ with respect to the change of $h$-th parameter from the value $p_h$ to the value $p_h'$ is calculated as in (Lecca et al., 2016) by

**FIGURE 1 |** The signalling of a DSB is initiated via the binding of the MRN complex which initiates resection. During HR, the ends of the double-strand break (DSB) are resected by nucleases, exposing single-strand DNA (ssDNA) that becomes bound by RPA. The mediator protein, BRCA2 initiates the loading of RAD51 onto ssDNA, helping to displace RPA. RAD51 oligomerizes, forms a nucleoprotein filament, and then searches for the homologous DNA sequence on the intact chromosome. RAD51 filament invades the intact dsDNA and forms a D-loop structure. It is further processed by DNA polymerases, chromatin remodelers (RAD54), nucleases, and ligases to restore it back to its original sequences. (Adapted from (Rossi et al., 2021).

**FIGURE 2 |** Conversion of the SIF format interactions into a (hyper-)graph structure.

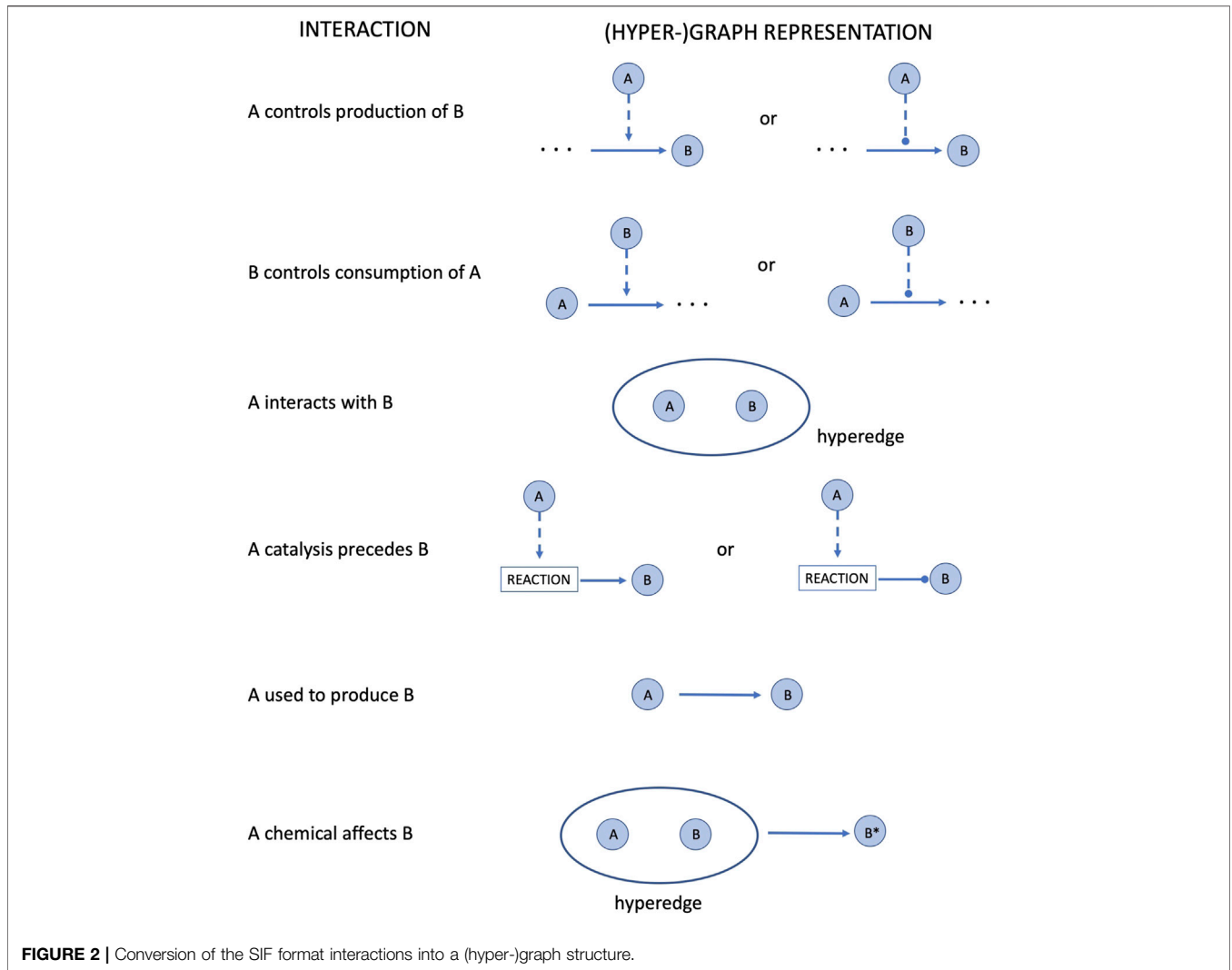$$SI_{ish} = \frac{1}{N} \sum_{k=1}^{N} \left( \frac{1}{N_P - 1} \sum_{r=1}^{N_P} \left( x_s^{(r)}(t_k | p_h \leftarrow p_h') - \overline{x_s}^{(r)}(p_h \leftarrow p_h') \right)^2 \right)^{\frac{1}{2}}$$

(1)

where $N$ is the length of the time series $x_s(t)$, and

$$\overline{x_s}(p_h \leftarrow p_h') = \frac{1}{N_P} \sum_{k=1}^{N_P} x_s(t_k | p_h \leftarrow p_h').$$

(2)

where "$p_h \leftarrow p_h'$" means "$p_h$ replaced by $p_h'$". The **Eq. 1** defines the mean of the standard deviations of the distributions of the simulated values of a protein/gene abundance at time points $t_k$.

In case the user knows the values of the rate constants, he/she can add them as an extra column to the SIF format of the input files. At the moment of writing, for most interactions the values of the kinetic constants are not known and there are no time-resolved data from which it is possible to infer them. It is precisely this context that justifies our choice to study the dynamics of the

**TABLE 3 |** Translation of BioPAX interactions into simple ordinary differential equations. See in **Figure 2** the (hyper-)graph representation of these interactions.

| Interaction | Differential equation |
|---|---|
| A controls production of B | $\frac{dB}{dt} = kA, \frac{dA}{dt} = 0$ |
| B controls consumption of A | $\frac{dA}{dt} = -kB, \frac{dB}{dt} = 0$ |
| A interacts with B | $\frac{dA}{dt} = \frac{dB}{dt} = -kAB$ |
| A catalysis precedes B | $\frac{dB}{dt} = \pm kA$ |
| A used to produce B | $\frac{dB}{dt} = kA$ |
| A chemical affects B | $\frac{dA}{dt} = 0, \frac{dB}{dt} = -kB = -\frac{dB^*}{dt}$ |

system in a range of values of the model parameters and more generally to provide a software that can be used as a platform for in silico experiments.

We implemented Network Analyser and Dynamics Simulator (NADS) consists of three modules written in R language:
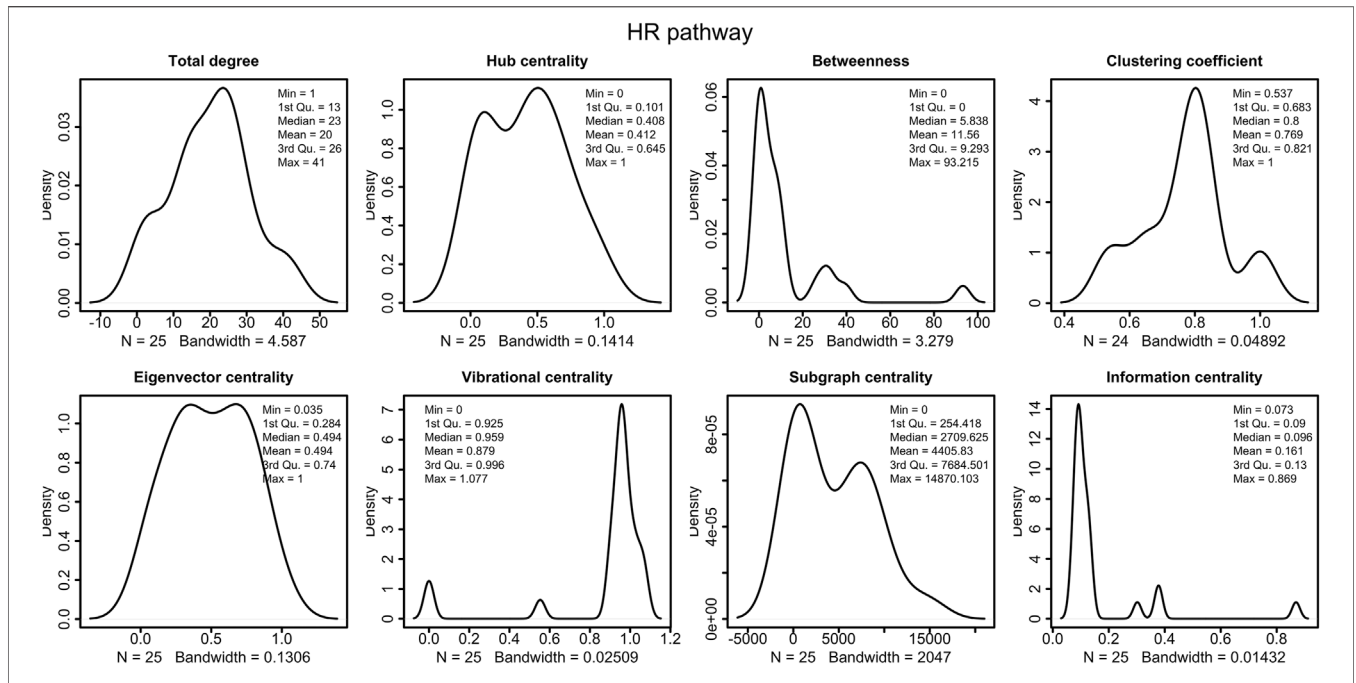
**FIGURE 3 |** Distributions of the centrality measures of HR pathway (Orlic-Milacic, 2015). We observe that the majority of node have low betweenness, low information centrality and high vibrational centrality.



**FIGURE 4 |** HR network and FANCM network. Colors vary from yellow to green according to increasing degree values. Node sizes grow as the betweenness centrality of nodes.

- network_analysis_functions.R: this module implements the functions that processes the SIF data-frames to make them suitable to their conversion into a graph. This module also implement the functions to rank the nodes according to their centrality measures;
- graph_parser.R: this module converts the SIF format network into an R script that solves the corresponding ordinary differential equations;

- network_analysis.R: this module first calls network_analysis_functions.R and performs the networks analysis, and secondly it calls the module graph_parser.R that generates the script dynamics.R containing the differential equations of the network dynamics.

The user can launch the software simply by running in RStudio the script network_analysis.R, and then by answering

**TABLE 4 |** Values of the centrality measures for the HR pathway in (Orlic-Milacic, 2015). In bold, we marked the genes/proteins with the highest scores.

| Protein | Total degree | In-degree | Out-degree | Hub centrality | Betweenness | Clustering coefficient | Eigenvector centrality | Vibrational centrality | Subgraph centrality | Information centrality |
|---|---|---|---|---|---|---|---|---|---|---|
| BLM | 30 | 19 | 11 | 4.08E-01 | 39.5138622 | 0.2183908 | 0.81043641 | **1.0771148** | 7408.13084 | 0.08351073 |
| EME1 | 13 | 0 | 13 | 5.21E-01 | 0 | 0.5384615 | 0.2589557 | 0.9579369 | 2709.62496 | 0.13185164 |
| MRE11 | 26 | 16 | 10 | 3.69E-01 | 9.2926175 | 0.1969231 | 0.73968044 | 1.0587524 | 6020.65076 | 0.08957011 |
| MUS81 | 13 | 1 | 12 | 5.18E-01 | 0 | 0.5384615 | 0.2589557 | 0.9547013 | 2291.00787 | 0.13177729 |
| NBN | 27 | 17 | 10 | 3.67E-01 | 9.2926175 | 0.1823362 | 0.76703983 | 1.0603542 | 6020.65076 | 0.08786049 |
| POLD1 | 26 | 9 | 17 | 6.83E-01 | 7.8121197 | 0.36 | 0.67253587 | 0.9655278 | 7684.50086 | 0.08965382 |
| POLD2 | 23 | 5 | 18 | 7.10E-01 | 5.8377525 | 0.4624506 | 0.57100865 | 0.9672104 | 9065.81813 | 0.09552092 |
| POLD3 | 23 | 8 | 15 | 6.45E-01 | 5.8377525 | 0.4624506 | 0.57100865 | 0.961071 | 5521.89505 | 0.09524672 |
| POLD4 | 23 | 7 | 16 | 6.66E-01 | 5.8377525 | 0.4624506 | 0.57100865 | 0.963361 | 6513.86378 | 0.09533952 |
| **RAD50** | 27 | 13 | 14 | 5.79E-01 | 10.2729789 | 0.2108262 | 0.75819149 | 1.0206621 | **11427.9652** | 0.08811688 |
| RAD51 | 23 | 12 | 11 | 4.03E-01 | 30.3997662 | 0.4426877 | 0.49447273 | 0.9038147 | 1403.03278 | 0.09426237 |
| RAD51B | 14 | 12 | 2 | 8.58E-02 | 0.3636364 | 0.7032967 | 0.31841389 | 0.9960094 | 74.47586 | 0.12453711 |
| RAD51C | 21 | 9 | 12 | 4.96E-01 | 1.6614219 | 0.4095238 | 0.47619405 | 0.9027594 | 1911.88011 | 0.09955242 |
| RAD51D | 16 | 11 | 5 | 2.15E-01 | 0.5127787 | 0.7166667 | 0.3593928 | 0.9248884 | 282.05774 | 0.11524626 |
| **RAD52** | 4 | 2 | 2 | 1.01E-01 | 0 | **1** | 0.1247068 | 0.9567526 | 254.41809 | 0.30199153 |
| **RPA1** | **41** | 20 | 21 | 9.02E-01 | **93.2152627** | 0.1512195 | 0.89284007 | 0.9069973 | 8617.68902 | 0.0726257 |
| **RPA2** | **41** | **22** | 19 | 8.46E-01 | 32.6964097 | 0.1487805 | **1** | 0.9404821 | 8508.98841 | 0.07306925 |
| **RPA3** | 34 | 11 | **23** | **1.00E+00** | 25.979693 | 0.2174688 | 0.819185 | 0.9803911 | **14870.10272** | 0.07928594 |
| **RPA4** | 3 | 0 | 3 | 1.27E-01 | 0 | **1** | 0.10547581 | 0.9592952 | 700.52299 | 0.37822514 |
| **SEM1** | 3 | 1 | 2 | 9.14E-02 | 0 | **1** | 0.06803348 | 0.9398974 | 63.99639 | 0.37757664 |
| TOP3A | 24 | 12 | 12 | 4.64E-01 | 9.2926175 | 0.2318841 | 0.68062612 | 1.0251952 | 8739.44955 | 0.09349123 |
| XRCC2 | 15 | 14 | 1 | 5.60E-02 | 0.8476272 | 0.7142857 | 0.34875334 | 1.0105203 | 55.0254 | 0.11916455 |
| XRCC3 | 13 | 12 | 1 | 4.68E-02 | 0.3333333 | 0.8076923 | 0.2842042 | 0.5530372 | 0 | 0.13016119 |
| BRCA2 | 16 | 16 | 0 | 1.40E-16 | 0 | 0.45 | 0.36198393 | 0 | 0 | 0.11465992 |
| **TOP3B** | 1 | 1 | 0 | 8.77E-18 | 0 | NA | 0.03472425 | 0 | 0 | **0.86898249** |

to the questions in the interactive interface as shown in **Table 2**.

We provide also the fourth module implementing the parametric sensitivity analysis, named sensitivity_analysis.R, which takes as an input the system of equations automatically generated by network_analysis.R.

The main module is the script network_analysis.R. As soon as the user runs it from the R Studio GUI (R Studio, 2022) or from a terminal, an interactive output is displayed as in **Tables 2**, **3**. The program asks the user to select the network to be analysed and then.

- it calculates the centralities measures
- it translates the SIF network into a hyper-graph structure according to the rules reported in **Figure 2**
- then, it translates the hyper-graph into a set of ordinary differential equations, according to the rules reported in **Table 3**,
- and, finally, it solves them.

The program returns also the execution times for the tasks expected to be the most computationally demanding, such as integrating the equation.

## 4 RESULTS

The HR network considered in this study has 25 nodes and 250 edges, as reported in **Supplementary Tables S1–S3**. The right part of the **Figure 4** shows the HR network in circular layout. The

network analysis phase of our study calculated the centrality measures distributions show in **Figure 3**, and identified six genes, as shown in **Table 4**:

1. BLM, scoring first for vibrational centrality
2. RAD50 scores first for sub-graph centrality
3. RAD52, scoring first for clustering coefficient
4. RPA1, scoring first for total degree and betweenness
5. RPA2, scoring first for in-degree and eigenvector centrality
6. RPA3, scoring first for out-degree, hub centrality, and sub-graph centrality
7. RPA4, scoring first for clustering coefficient
8. SEM1, scoring first for clustering coefficient.

Of particular interest is the fact that BLM has the highest vibration centre. The interpretation of this result is that BLM is the node most sensitive to stresses and/or stimuli, i.e., according to the vibrational centrality measure, it is the most vulnerable node in the network (Estrada and Hatano, 2010). This result is of particular interest in light of the crucial role this gene plays in the HR network. Indeed, the key role of BLM is well know, and alterations in this protein is linked to different diseases including cancer (Kaur et al., 2021). BLM is a $3'$-$5'$ ATP-dependent RecQ DNA helicase. It is a genome stabilizer playing an essential role in the DNA replication regulation, DNA recombination, and both homologous and non-homologous pathways of DSB repair. The high vulnerability of the BLM node to external stimuli and conditions suggests the need to identify which conditions and/or stimuli may be altering it, in order to preserve its proper functioning and/or to understand how it can be restored if it is

**TABLE 5** | This is the PART I of the table of ordinary differential equations of the dynamics of HR network, in R code formalism. The *k* followed by a number denote the kinetic rate constant, and the letter "d" in front of the name of the proteins denote the temporal derivative of it concentration.

```
dBLM   =  - k12*EME1  - k34*MRE11*BLM  - k35*MUS81  - k55*NBN*BLM
  - k57*POLD1  - k74*POLD2  - k92*POLD3  - k107*POLD4  - k132*RAD50*BLM
  - k138*RAD51*BLM  - k162*RAD51D*BLM  - k169*RPA1  - k182*RPA1*BLM
  - k190*RPA2  - k202*RPA2*BLM  - k209*RPA3  - k222*RPA3*BLM
  - k248*TOP3A*BLM  - k249*XRCC2*BLM

dBRCA2  =  - k13*EME1 - k36*MUS81 - k58*POLD1 - k75*POLD2 - k93*POLD3
  - k108*POLD4  - k139*RAD51*BRCA2  - k150*RAD51C
    - k155*RAD51C*BRCA2  - k170*RPA1  - k183*RPA1*BRCA2
  - k191*RPA2  - k203*RPA2*BRCA2  - k210*RPA3  - k235*SEM1*BRCA2
  - k250*XRCC3*BRCA2

dMRE11  =  - k1*BLM  - k14*EME1  - k34*MRE11*BLM  - k37*MUS81
  - k56*NBN*MRE11  - k59*POLD1  - k76*POLD2  - k94*POLD3
  - k109*POLD4  - k133*RAD50*MRE11  - k171*RPA1  - k184*RPA1*MRE11
  - k192*RPA2  - k204*RPA2*MRE11  - k211*RPA3  - k223*RPA3*MRE11
  - k237*TOP3A

dNBN   =  - k2*BLM  - k15*EME1  - k32*MRE11  - k38*MUS81  - k55*NBN*BLM
  - k56*NBN*MRE11  - k60*POLD1  - k77*POLD2  - k95*POLD3  - k110*POLD4
  - k130*RAD50  - k134*RAD50*NBN  - k172*RPA1  - k185*RPA1*NBN
  - k193*RPA2  - k205*RPA2*NBN  - k212*RPA3  - k224*RPA3*NBN
  - k238*TOP3A

dPOLD1  =  - k3*BLM  - k25*MRE11  - k47*NBN  - k72*POLD1*POLD3
  - k73*POLD1*POLD4  - k89*POLD2*POLD1  - k123*RAD50  - k186*RPA1*POLD1
  - k206*RPA2*POLD1  - k225*RPA3*POLD1  - k239*TOP3A

dPOLD2  =  - k4*BLM  - k26*MRE11  - k48*NBN  - k89*POLD2*POLD1
  - k90*POLD2*POLD3  - k91*POLD2*POLD4  - k124*RAD50  - k240*TOP3A

dPOLD3  =  - k5*BLM  - k27*MRE11  - k49*NBN  - k72*POLD1*POLD3
  - k90*POLD2*POLD3  - k122*POLD4*POLD3  - k125*RAD50  - k241*TOP3A

dPOLD4  =  - k6*BLM  - k28*MRE11  - k50*NBN  - k73*POLD1*POLD4
  - k91*POLD2*POLD4  - k122*POLD4*POLD3  - k126*RAD50  - k242*TOP3A

dRAD50  =  - k7*BLM  - k16*EME1  - k39*MUS81  - k61*POLD1  - k78*POLD2
  - k96*POLD3  - k111*POLD4  - k132*RAD50*BLM  - k133*RAD50*MRE11
  - k134*RAD50*NBN  - k135*RAD50*RPA1  - k136*RAD50*RPA2  - k137*RAD51
  - k173*RPA1  - k194*RPA2  - k213*RPA3  - k226*RPA3*RAD50  - k243*TOP3A

dRAD51  =  - k17*EME1  - k40*MUS81  - k62*POLD1  - k79*POLD2  - k97*POLD3
  - k112*POLD4  - k138*RAD51*BLM  - k139*RAD51*BRCA2  - k140*RAD51*RAD51B
  - k141*RAD51*RAD51D  - k142*RAD51*RAD52  - k143*RAD51*RPA1  - k144*RAD51*RPA2
  - k145*RAD51*SEM1  - k146*RAD51*XRCC2  - k147*RAD51*XRCC3  - k151*RAD51C
  - k156*RAD51C*RAD51  - k174*RPA1  - k195*RPA2  - k214*RPA3  - k227*RPA3*RAD51

dRAD51B  =  - k18*EME1  - k41*MUS81  - k63*POLD1  - k80*POLD2  - k98*POLD3
  - k113*POLD4  - k140*RAD51*RAD51B  - k148*RAD51B*RPA1  - k149*RAD51B*XRCC2
  - k157*RAD51C*RAD51B  - k163*RAD51D*RAD51B  - k175*RPA1  - k196*RPA2
  - k215*RPA3

dRAD51C  =  - k19*EME1  - k42*MUS81  - k64*POLD1  - k81*POLD2  - k99*POLD3
  - k114*POLD4  - k155*RAD51C*BRCA2  - k156*RAD51C*RAD51  - k157*RAD51C*RAD51B
  - k158*RAD51C*RAD51D  - k159*RAD51C*RPA1  - k160*RAD51C*XRCC2
  - k161*RAD51C*XRCC3  - k176*RPA1  - k197*RPA2  - k216*RPA3
```

**TABLE 6 |** This is the PART II (continuation) of the table of ordinary differential equations of the dynamics of HR network, in R code formalism. The *k* followed by a number denote the kinetic rate constant, and the letter "d" in front of the name of the proteins denote the temporal derivative of it concentration.

```
dRAD51D  =  − k20*EME1 − k43*MUS81  − k65*POLD1  − k82*POLD2
  − k100*POLD3 − k115*POLD4  − k141*RAD51*RAD51D  − k158*RAD51C*RAD51D
  − k162*RAD51D*BLM  − k163*RAD51D*RAD51B  − k164*RAD51D*RPA1
  − k165*RAD51D*XRCC2 − k166*RAD51D*XRCC3  − k177*RPA1
  − k198*RPA2  − k217*RPA3

dRAD52  =  −k142*RAD51*RAD52  − k167*RAD52*RPA1
  − k168*RAD52*RPA2  − k228*RPA3*RAD52

dRPA1  =  − k8*BLM − k29*MRE11  − k51*NBN  − k66*POLD1  − k83*POLD2
  − k101*POLD3  − k116*POLD4  − k127*RAD50  − k135*RAD50*RPA1
  − k143*RAD51*RPA1 − k148*RAD51B*RPA1  − k152*RAD51C  − k159*RAD51C*RPA1
  − k164*RAD51D*RPA1 − k167*RAD52*RPA1  − k182*RPA1*BLM  − k183*RPA1*BRCA2
  − k184*RPA1*MRE11 − k185*RPA1*NBN  − k186*RPA1*POLD1  − k187*RPA1*TOP3A
  − k188*RPA1*TOP3B − k189*RPA1*XRCC2  − k207*RPA2*RPA1  − k229*RPA3*RPA1
  − k232*RPA4*RPA1 − k236*SEM1*RPA1  − k244*TOP3A

dRPA2  =  − k9*BLM  − k11*BLM  − k30*MRE11  − k33*MRE11  − k52*NBN
  − k54*NBN − k67*POLD1  − k84*POLD2  − k102*POLD3  − k117*POLD4
  − k128*RAD50  − k131*RAD50 − k136*RAD50*RPA2  − k144*RAD51*RPA2
  − k153*RAD51C  − k168*RAD52*RPA2 − k178*RPA1  − k202*RPA2*BLM
  − k203*RPA2*BRCA2  − k204*RPA2*MRE11 − k205*RPA2*NBN  − k206*RPA2*POLD1
  − k207*RPA2*RPA1  − k208*RPA2*TOP3A − k218*RPA3  − k230*RPA3*RPA2
  − k233*RPA4*RPA2  − k245*TOP3A  − k247*TOP3A

dRPA3  =  − k10*BLM  − k31*MRE11  − k53*NBN  − k68*POLD1  − k85*POLD2
  − k103*POLD3  − k118*POLD4  − k129*RAD50  − k154*RAD51C  − k222*RPA3*BLM
  − k223*RPA3*MRE11  − k224*RPA3*NBN  − k225*RPA3*POLD1  − k226*RPA3*RAD50
  − k227*RPA3*RAD51  − k228*RPA3*RAD52  − k229*RPA3*RPA1  − k230*RPA3*RPA2
  − k231*RPA3*TOP3A  − k234*RPA4*RPA3  − k246*TOP3A

dRPA4  =  −k232*RPA4*RPA1  − k233*RPA4*RPA2  − k234*RPA4*RPA3

dSEM1  =  −k145*RAD51*SEM1  − k235*SEM1*BRCA2  − k236*SEM1*RPA1

dTOP3A  =  − k21*EME1  − k44*MUS81  − k69*POLD1  − k86*POLD2
  − k104*POLD3 − k119*POLD4  − k179*RPA1  − k187*RPA1*TOP3A  − k199*RPA2
  − k208*RPA2*TOP3A  − k219*RPA3  − k231*RPA3*TOP3A  − k248*TOP3A*BLM


dEME1  =  −k24*EME1*MUS81

dMUS81  =  −k24*EME1*MUS81

dTOP3B  =  −k188*RPA1*TOP3B
```

altered. The high vulnerability of this node could also be explained by a recent study by Kaur et al. (2021). These authors report that BLM has a dual function both as a tumour suppressor and possibly as a proto-oncogene, being probably involved in the mechanisms of its deregulation in tumours.

The analysis also correctly identifies the SEM1 gene as a node with a high clustering coefficient. Indeed, as reported in (Safran et al., 2021; GeneCard, 2022), SEM1 gene encodes for a protein that is part of a 26S proteasome, which is a multiprotein complex with a function in the ATP-dependent degradation of ubiquitinated proteins. This complex contributes to the maintenance of protein homeostasis by removing misfolded or damaged proteins, which could jeopardize the healthy cellular functions, and by removing proteins no longer need. Therefore, 26S proteasome is involved in numerous cellular processes, including cell cycle progression, apoptosis, or DNA damage repair (Sone et al., 2004).

**TABLE 7 |** For each gene/protein in the HR pathway in (Orlic-Milacic, 2015) we selected the kinetic rates whose sensitivity index belongs to the 98th percentile of the sensitivity index distribution. The sensitivity index is calculated using the **formula (1)**.

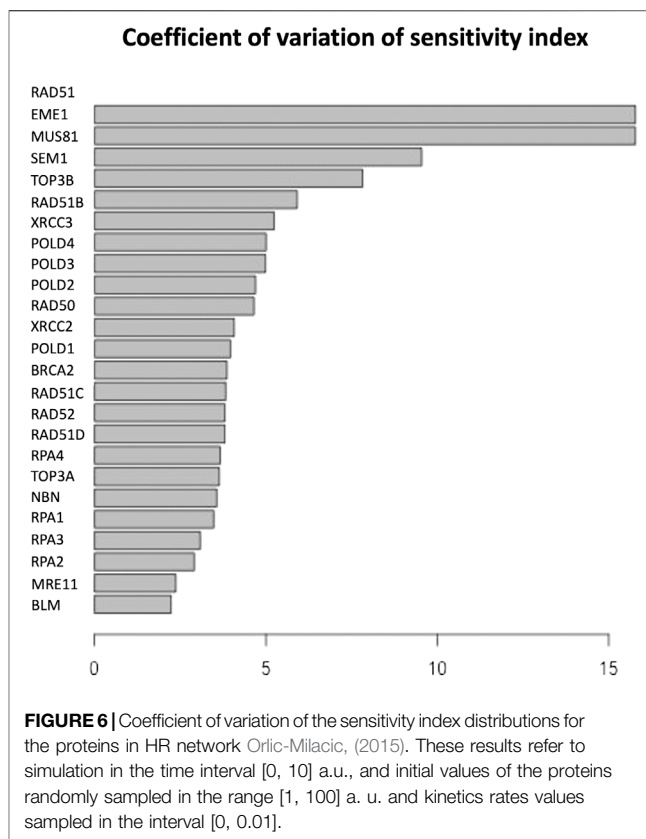| Gene | Kinetic rate to which it is highly sensitive |
|------|----------------------------------------------|
| BLM | k12, k34, k107, k138, k222 |
| BRCA2 | k13, k58, k108, k139, k250 |
| EME1 | k24, k138, k140, k144, k227 |
| MRE11 | k14, k34, k109, k204, k223 |
| MUS81 | k24, k138, k140, k144, k227 |
| NBN | k15, k32, k34, k60, k110 |
| POLD1 | k72, k73, k89, k90, k91 |
| POLD2 | k72, k73, k89, k90, k91 |
| POLD3 | k72, k89, k90, k91, k122 |
| POLD4 | k73, k89, k90, k91, k122 |
| RAD50 | k16, k61, k78, k111, k137 |
| RAD51 | – |
| RAD51B | k18, k63, k113, k140, k163 |
| RAD51C | k19, k64, k81, k114, k157 |
| RAD51D | k20, k65, k115, k141, k162 |
| RAD52 | k142, k168, k222, k227, k228 |
| RPA1 | k8, k29, k66, k83, k116 |
| RPA2 | k30, k33, k117, k144, k206 |
| RPA3 | k31, k68, k118, k222, k227 |
| RPA4 | k29, k227, k232, k233, k234 |
| SEM1 | k139, k145, k235, k236, k250 |
| TOP3A | k21, k69, k86, k119, k248 |
| TOP3B | k29, k66, k83, k116, k188 |
| XRCC2 | k22, k70, k120, k146, k149 |
| XRCC3 | k23, k121, k147, k166, k250 |



**FIGURE 6 |** Coefficient of variation of the sensitivity index distributions for the proteins in HR network Orlic-Milacic, (2015). These results refer to simulation in the time interval [0, 10] a.u., and initial values of the proteins randomly sampled in the range [1, 100] a. u. and kinetics rates values sampled in the interval [0, 0.01].



**FIGURE 5 |** Mean of the sensitivity index distributions for the proteins in HR network (Orlic-Milacic, 2015). These results refer to simulation in the time interval [0, 10] a.u., and initial values of the proteins randomly sampled in the range [1, 100] a. u. and kinetics rates values sampled in the interval [0, 0.01].
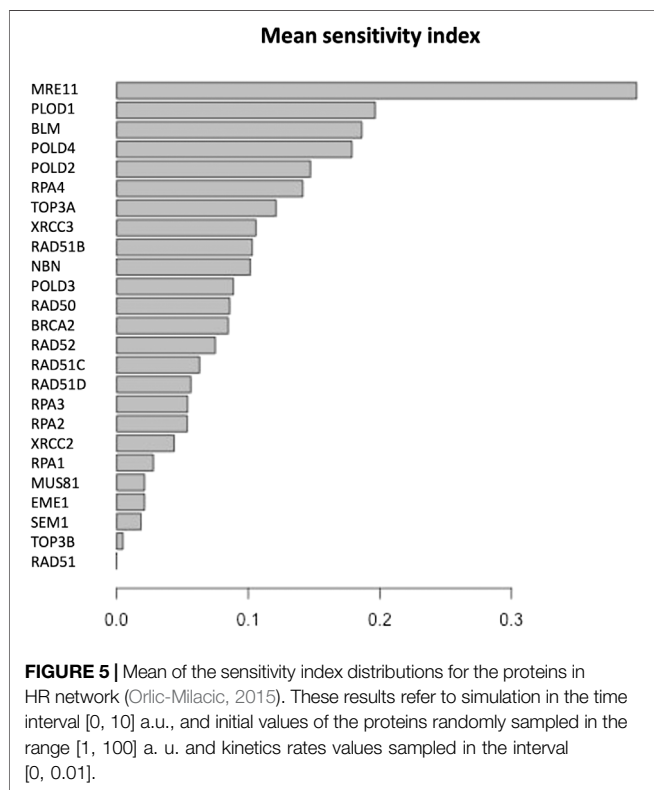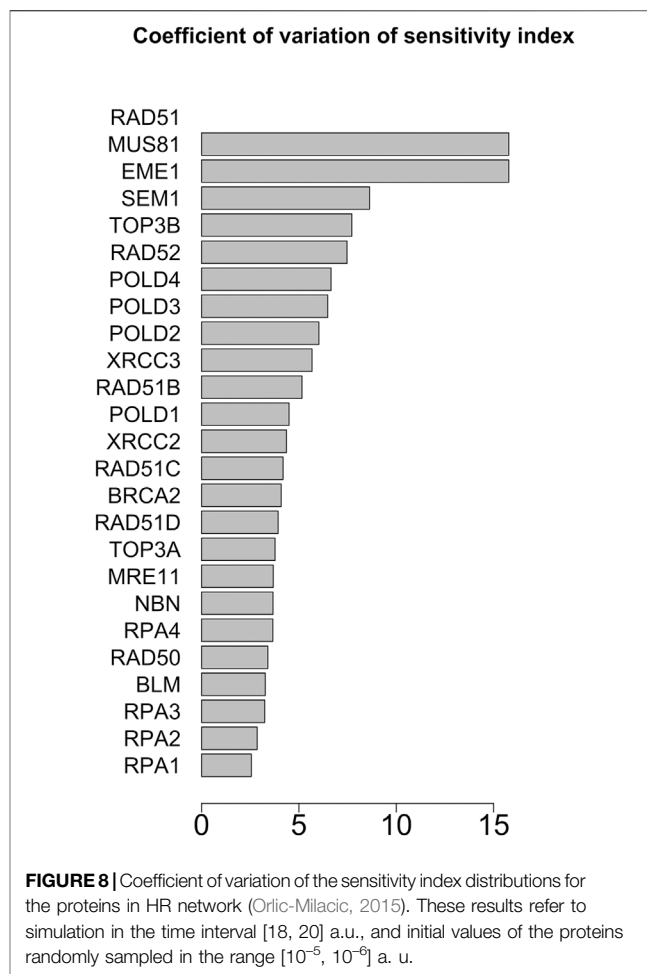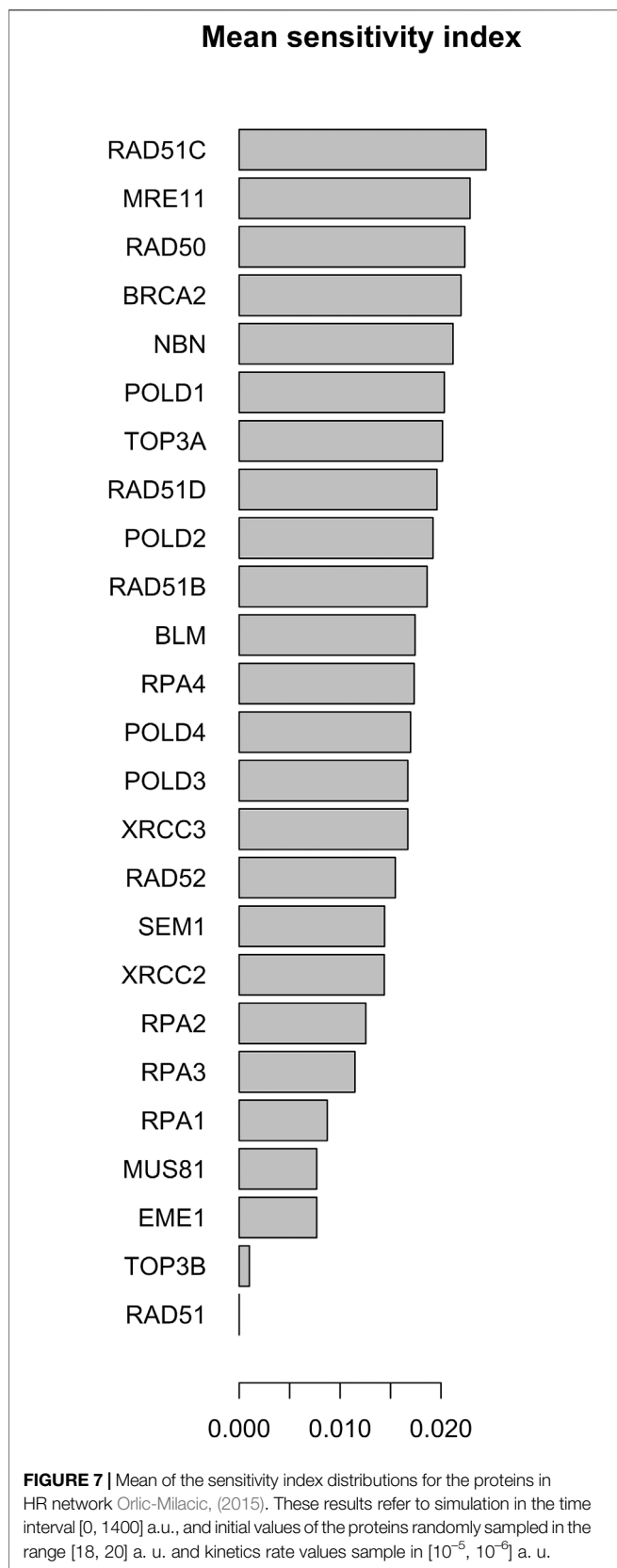
SEM1 was found also as a subunit in experiments of affinity purification of the yeast 19S proteasome, and its human homolog, DSS1, was found to copurify with the human 19S proteasome (Krogan et al., 2004). DSS1 is associated with the tumour suppressor protein BRCA2 involved in DNA DSBs repair. The authors in (Krogan et al., 2004) proved that SEM1 is essential for efficient repair of an HO-generated yeast DSB using both HR and nonhomologous end joining (NHEJ) pathways. Moreover, they showed that deletion of SEM1 contributes to cause defects in (synthetic) growth and hypersensitivity to genotoxins when combined with mutations in certain well-established genes involved in the DNA DSB repair.

Similarly to SEM1, the result of a high clustering coefficient is also expected for RPA4, as RPA4 is also part of a complex (Keshav et al., 1995). RPA4 gene encodes a single-stranded DNA-binding protein that is a subunit of the replication protein A complex (GeneCard, 2022). Replication protein A is essential for DNA DSB repair and plays a crucial role in the activation of cell cycle checkpoints. As regards the RPA complex, we have already seen in the previous sections that the RPA complex controls DNA repair and DNA damage checkpoint activation as well. In particular, the network analysis shows that RPA1 highly scores by total degree and betweenness. These results reflect the fact that RPA1 is an active route of communication exchanges between various nodes in the network. RPA1 is part of the heterotrimeric replication protein A complex (RPA/RP-A). It stabilizes single-

## Mean sensitivity index



**FIGURE 7 |** Mean of the sensitivity index distributions for the proteins in HR network Orlic-Milacic, (2015). These results refer to simulation in the time interval [0, 1400] a.u., and initial values of the proteins randomly sampled in the range [18, 20] a. u. and kinetics rate values sample in [10⁻⁵, 10⁻⁶] a. u.

## Coefficient of variation of sensitivity index



**FIGURE 8 |** Coefficient of variation of the sensitivity index distributions for the proteins in HR network (Orlic-Milacic, 2015). These results refer to simulation in the time interval [18, 20] a.u., and initial values of the proteins randomly sampled in the range [10⁻⁵, 10⁻⁶] a. u.

stranded DNA intermediates, that form during DNA replication or upon DNA stress In (Bass et al., 2016; Haahr et al., 2016; Human Protein Atlas, 2022). It prevents the reannealing of single-stranded DNA intermediates and recruits and activates different proteins and complexes forming part of DNA metabolism. Thereby, it is a key protein both in DNA replication and in the cellular response to DNA damage (Lin et al., 1998). RPA2 shows high score for in-degree and eigenvevtor centrality, meaning that it is interacting with protein also highly scoring by eigenvector centrality and degree (Hansen et al., 2020), and thence with proteins which have a great influence in the HR network. Indeed RPA 2 gene has been found highly expressed in low grade carcinomas and its expression has a gradual significant decrease from stage I to stage IV carcinomas. All the three subunits RPA1, RPA2, and RPA3, were more abundant (with statistical significance evidence) in lymph node negative and earlier stage (stage I and II) gastric carcinomas (Fourtziala et al., 2020). Finally, of particular interest and the fact that RPA3 ranks first in terms of centrality out-degree, hub-centrality and sub-graph centrality. Since subgraph centrality of a node is the number of subgraphs a node
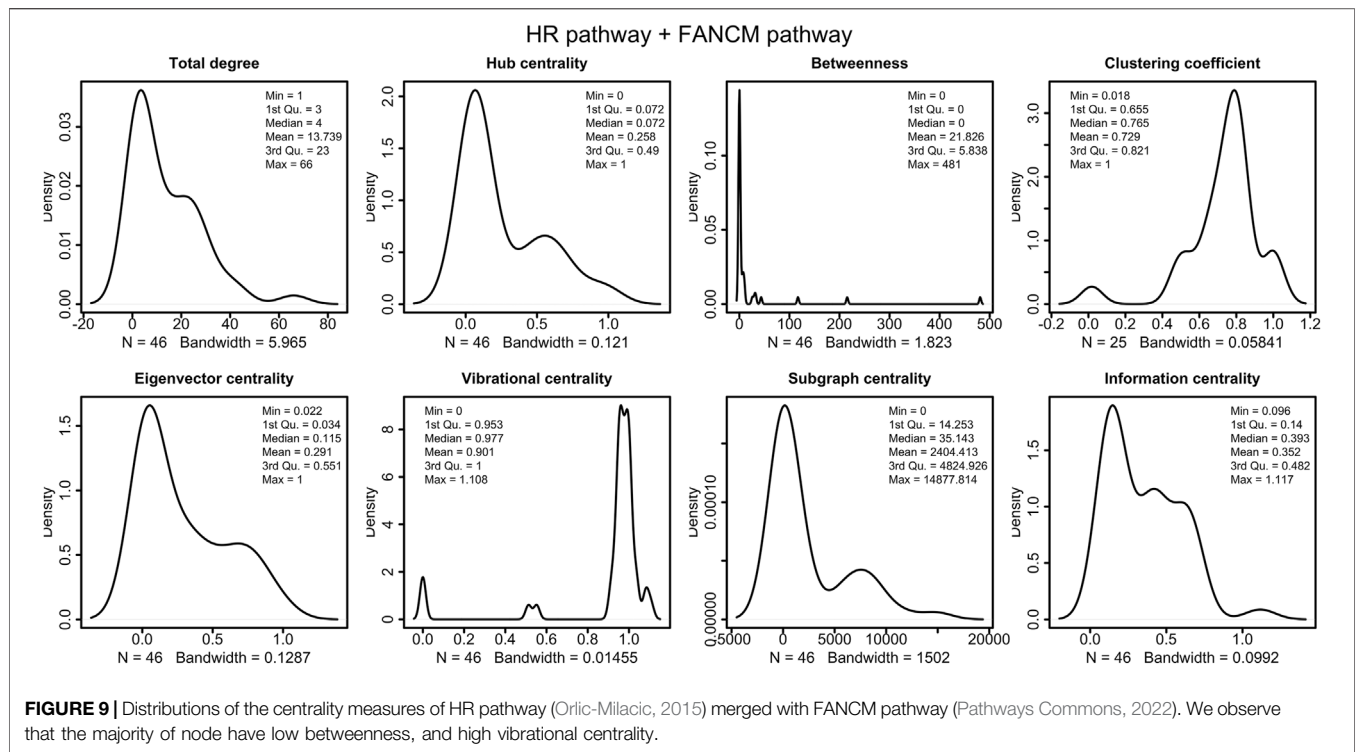
**TABLE 8** | Values of the centrality measures for the HR pathway in (Orlic-Milacic, 2015) merged with the FANCM pathway in (Pathways Commons, 2022). In bold, we marked the genes with the highest scores.

| Protein | Total degree | In-degree | Out-degree | Hub centrality | Betweenness | Clustering coefficient | Eigenvector centrality | Vibrational centrality | Subgraph centrality | Information centrality |
|---|---|---|---|---|---|---|---|---|---|---|
| BLM | 32 | 19 | 13 | 4.83E-01 | 117.0805289 | 0.72058824 | 0.83056114 | **1.107752** | 7408.39241 | 0.1204394 |
| EME1 | 15 | 0 | 15 | 5.92E-01 | 0 | 0.48351648 | 0.28101096 | 9.68E-01 | 2727.280253 | 0.1730775 |
| MRE11 | 26 | 16 | 10 | 3.71E-01 | 9.2926175 | 0.82051282 | 0.73943396 | 1.08E+00 | 6025.93013 | 0.1319644 |
| MUS81 | 13 | 1 | 12 | 5.18E-01 | 0 | 0.53846154 | 0.25936455 | 9.56E-01 | 2292.062046 | 0.1868928 |
| NBN | 27 | 17 | 10 | 3.69E-01 | 9.2926175 | 0.82051282 | 0.76658977 | 1.08E+00 | 6025.93013 | 0.1296602 |
| POLD1 | 26 | 9 | 17 | 6.82E-01 | 7.8121197 | 0.76470588 | 0.67199489 | 9.86E-01 | 7688.630996 | 0.1322165 |
| POLD2 | 23 | 5 | 18 | 7.10E-01 | 5.8377525 | 0.76470588 | 0.57047643 | 9.81E-01 | 9071.417868 | 0.1400556 |
| POLD3 | 23 | 8 | 15 | 6.44E-01 | 5.8377525 | 0.76470588 | 0.57047643 | 9.78E-01 | 5524.141817 | 0.1396592 |
| POLD4 | 23 | 7 | 16 | 6.65E-01 | 5.8377525 | 0.76470588 | 0.57047643 | 9.80E-01 | 6516.909994 | 0.1397968 |
| **RAD50** | 27 | 13 | 14 | 5.81E-01 | 10.2729789 | 0.81318681 | 0.75778635 | 1.04E+00 | **11437.382893** | 0.1300547 |
| RAD51 | 23 | 12 | 11 | 4.02E-01 | 30.3997662 | 0.65497076 | 0.4937734 | 9.17E-01 | 1403.106256 | 0.1386277 |
| RAD51B | 14 | 12 | 2 | 8.53E-02 | 0.3636364 | 0.82051282 | 0.31833841 | 1.01E+00 | 74.55065 | 0.1774293 |
| RAD51C | 21 | 9 | 12 | 4.93E-01 | 1.6614219 | 0.81904762 | 0.47562207 | 9.12E-01 | 1911.911354 | 0.1453357 |
| RAD51D | 16 | 11 | 5 | 2.13E-01 | 0.5127787 | 0.81904762 | 0.3597479 | 9.38E-01 | 282.219746 | 0.1656386 |
| RAD52 | 4 | 2 | 2 | 1.00E-01 | 0 | **1** | 0.12449815 | 9.63E-01 | 254.418337 | 0.3917785 |
| **RPA1** | 43 | 20 | **23** | 9.74E-01 | 215.4152627 | 0.49802372 | 0.91307865 | 9.47E-01 | 8620.183617 | 0.1074375 |
| **RPA2** | 41 | 22 | 19 | 8.45E-01 | 32.6964097 | 0.64210526 | **1** | 9.82E-01 | 8511.793141 | 0.1095086 |
| **RPA3** | 34 | 11 | 23 | **1** | 25.979693 | 0.64210526 | 0.81828849 | 9.98E-01 | 14877.81448 | 0.1182323 |
| **RPA4** | 3 | 0 | 3 | 1.27E-01 | 0 | **1** | 0.10543731 | 9.63E-01 | 700.745966 | 0.4815787 |
| **SEM1** | 3 | 1 | 2 | 9.07E-02 | 0 | **1** | 0.0682632 | 9.28E-01 | 64.063815 | 0.4810717 |
| **TOP3A** | 26 | 12 | 14 | 5.39E-01 | 43.5259509 | 0.73626374 | 0.70171805 | 1.04E+00 | 8739.804106 | 0.1322725 |
| XRCC2 | 15 | 14 | 1 | 5.59E-02 | 0.8476272 | 0.82417582 | 0.3492084 | 1.03E+00 | 55.032467 | 0.1704502 |
| XRCC3 | 13 | 12 | 1 | 4.65E-02 | 0.3333333 | 0.80769231 | 0.28356645 | 5.52E-01 | 0 | 0.1844967 |
| BRCA1 | 2 | 0 | 2 | 7.17E-02 | 0 | NA | 0.02248201 | 9.53E-01 | 15.252907 | 0.6489791 |
| CENPS | 2 | 0 | 2 | 7.17E-02 | 0 | NA | 0.02248201 | 9.53E-01 | 15.252907 | 0.6489791 |
| CENPX | 2 | 0 | 2 | 7.17E-02 | 0 | NA | 0.02248201 | 9.53E-01 | 15.252907 | 0.6489791 |
| EME2 | 2 | 0 | 2 | 7.17E-02 | 0 | NA | 0.02248201 | 9.53E-01 | 15.252907 | 0.6489791 |
| FAAP100 | 4 | 2 | 2 | 7.17E-02 | 0 | NA | 0.04496402 | 1.00E+00 | 14.252907 | 0.3943335 |
| FAAP20 | 2 | 0 | 2 | 7.17E-02 | 0 | NA | 0.02248201 | 9.53E-01 | 15.252907 | 0.6489791 |
| FAAP24 | 2 | 0 | 2 | 7.17E-02 | 0 | NA | 0.02248201 | 9.53E-01 | 15.252907 | 0.6489791 |
| FANCA | 4 | 2 | 2 | 7.17E-02 | 0 | NA | 0.04496402 | 1.00E+00 | 14.252907 | 0.3943335 |
| FANCB | 4 | 2 | 2 | 7.17E-02 | 0 | NA | 0.04496402 | 1.00E+00 | 14.252907 | 0.3943335 |
| FANCC | 4 | 2 | 2 | 7.17E-02 | 0 | NA | 0.04496402 | 1.00E+00 | 14.252907 | 0.3943335 |
| FANCD2 | 2 | 1 | 1 | 3.59E-02 | 0 | NA | 0.02248201 | 9.77E-01 | 3.563227 | 0.6504898 |
| FANCE | 4 | 2 | 2 | 7.17E-02 | 0 | NA | 0.04496402 | 1.00E+00 | 14.252907 | 0.3943335 |
| FANCF | 4 | 2 | 2 | 7.17E-02 | 0 | NA | 0.04496402 | 1.00E+00 | 14.252907 | 0.3943335 |
| FANCG | 4 | 2 | 2 | 7.17E-02 | 0 | NA | 0.04496402 | 1.00E+00 | 14.252907 | 0.3943335 |
| FANCI | 3 | 1 | 2 | 7.17E-02 | 0 | NA | 0.03372301 | 9.77E-01 | 14.252907 | 0.4827827 |
| FANCL | 4 | 2 | 2 | 7.17E-02 | 0 | NA | 0.04496402 | 1.00E+00 | 14.252907 | 0.3943335 |
| **FANCM** | **66** | **44** | 22 | 4.35E-16 | **481** | 0.01811594 | 0.29119966 | 5.13E-01 | 131.839388 | 0.0955692 |
| RMI1 | 2 | 0 | 2 | 7.17E-02 | 0 | NA | 0.02248201 | 9.53E-01 | 15.252907 | 0.6489791 |
| RMI2 | 2 | 0 | 2 | 7.17E-02 | 0 | NA | 0.02248201 | 9.53E-01 | 15.252907 | 0.6489791 |
| UBE2T | 3 | 2 | 1 | 3.59E-02 | 0 | NA | 0.03372301 | 1.00E+00 | 3.563227 | 0.4815848 |
| BRCA2 | 16 | 16 | 0 | 1.09E-16 | 0 | 0.69230769 | 0.36151501 | -4.08E-18 | 0 | 0.1645300 |
| TOP3B | 1 | 1 | 0 | 6.80E-18 | 0 | NA | 0.03524702 | 0.00E+00 | 0 | **1.1169817** |
| HES1 | 2 | 2 | 0 | 1.36E-17 | 0 | NA | 0.02248201 | 0.00E+00 | 0 | 0.6438225 |
| MAX SCORE | 43 | 22 | 23 | 0.9743705 | 215.4152627 | 0.82417582 | 0.91307865 | 1.08361 | 14877.81448 | 0 |

participates in (weighted according to their size) (Estrada and Rodríguez-Velázquez, 2005), it means that RPA3 take part into a number of subgraphs of significant size relatively to the whole network size. From our analysis it results that RPA3 versus RPA1 and RPA2, although its roles are similar to those of RPA2, has thus a great influence on pathways of crucial importance more than on subset of unconnected nodes or single nodes.

The analysis also highlights RAD50 that is a component of the MRN complex The protein complex is involved in numerous enzymatic activities required for nonhomologous joining of DNA ends. It is protein is essential for DNA double-strand break repair, cell cycle checkpoint activation, telomere maintenance, and meiotic recombination. (Carney et al., 1998; de Jager et al., 2001; Estrada and Ross, 2018; Bian et al., 2019; Beikzadeh and Latham, 2021; National library of Medicine, 2022). This role is reflected by th ehigh sungraph centrality that measures the centrality of a node by taking into account the number of subgraphs the node participates in. Specifically, the subgraph centrality of a node is the number of closed loops originating at

**FIGURE 9** | Distributions of the centrality measures of HR pathway (Orlic-Milacic, 2015) merged with FANCM pathway (Pathways Commons, 2022). We observe that the majority of node have low betweenness, and high vibrational centrality.

the node, where longer loops are exponentially downweighted. Consequently, subgraph centrality measures how close a node is to the other nodes in the network.

The results of this analysis reveal a correspondence between the measure of centrality and the role of the protein. On the basis of this, when the role of the protein is known, this information can be used to work out the correctness of the computational analysis. When, on the other hand, the role of the protein is not known, knowledge of its centrality measurements can suggest the type or set of types of possible roles.
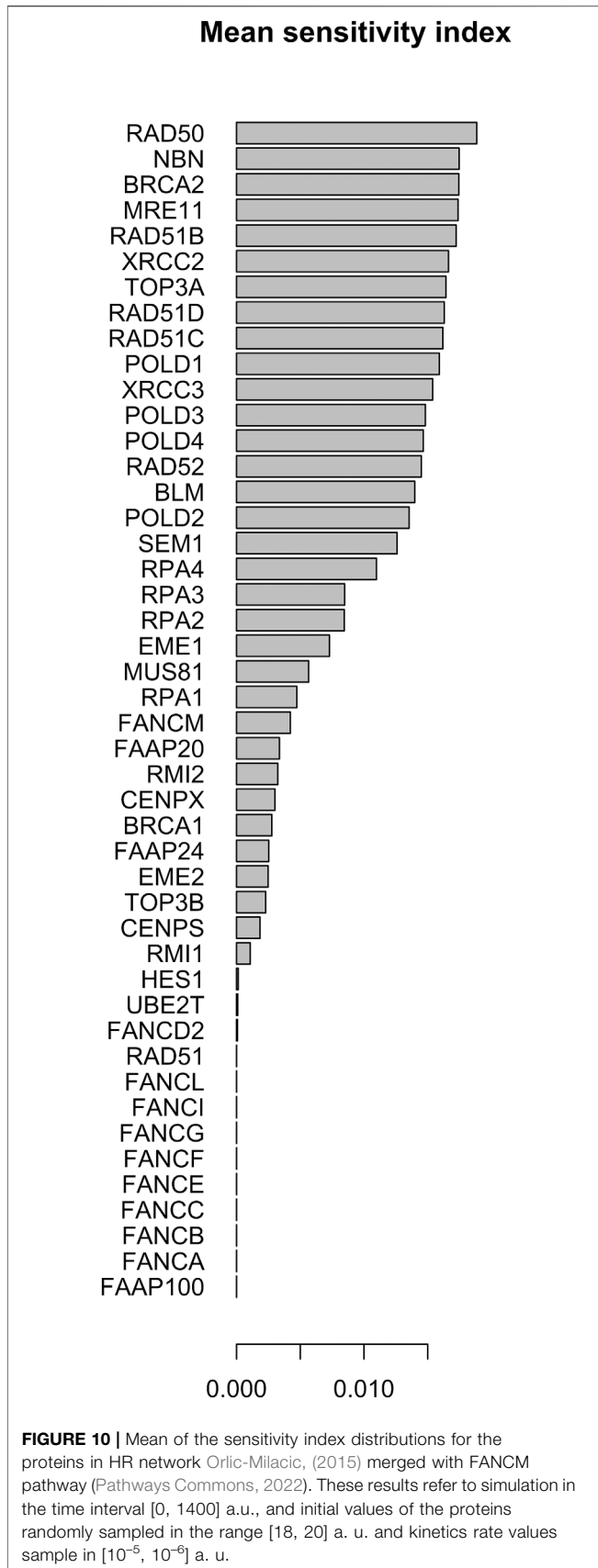
In **Tables 5**, **6** we show the rate equations of the HR network dynamics generated as the automatic translation of the network (see the script dynamics.R in the GitLab repository of NADS software). In file Simulations_of_Dynamics_HR_pathway.pdf provided in the Supplementary Material, we show the time evolution curves of each node of the HR network obtained as a solution of the equations.

The parameter sensitivity analysis was conducted by perturbing each parameter in the convergence range of the solution and yielded the results shown in **Table 7**; **Figures 5**, **6**. In **Table 7** we report the kinetic rate constants for which the sensitivity index belongs ot the 98th percentile of the sensitivity index distribution. They correspond to the most sensitive parameters, i.e. to the interactions whose alterations can significantly alter the dynamics of the network. To find out which interactions they refer to, the reader can refer to the **Supplementary Tables S1–S3**. **Figure 5** we shows that MRE11, followed by POLD1, BLM has the highest average sensitivity index. In **Figure 6** we show the coefficient of variation of each protein in the HR network. The coefficient of variation, being the ratio of the standard deviation to the mean,
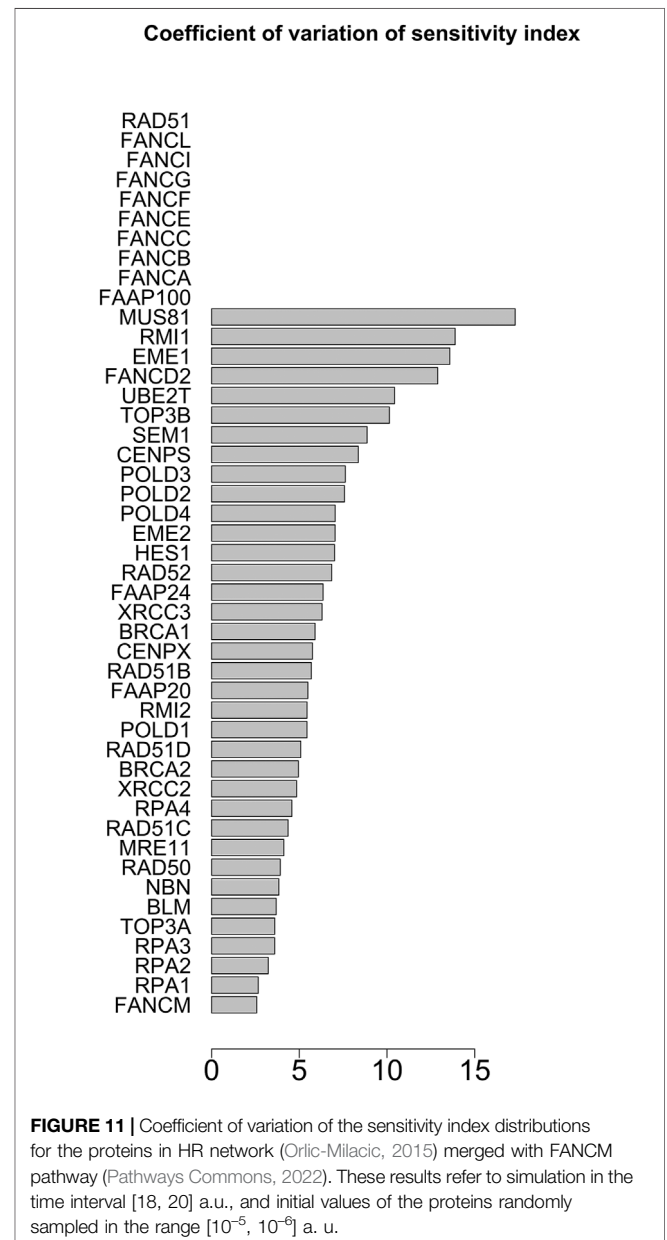
measures the extent of variability in relation to the mean of the population. The higher the coefficient of variation, the greater the dispersion. BLM, followed by MRE11 and RPA2 exhibits the lowest coefficient of variation of the sensitivity index. This mean that BLM, MRE11 and RPA2 have high sensitivity indices, and that the distribution of the sensitivity indices is well shaped around its mean, i.e. these protein exhibit almost the same sensitivity for all the parameters of the model. In this study we have therefore found that the BLM and RPA2 are sensitive nodes and that their sensitivity has two components: a topological sensitivity expressed by vibrational centrality, eigenvector centrality and clustering coefficient, and a dynamic sensitivity expressed by the parameter sensitivity index.

As with the BLM and RPA2 proteins, sensitivity analysis also highlights the MRE11 protein, which is highly sensitive to kinetic parameters, and its vibrational centrality is 1.0587524, very close to the maximum value exhibited by BLM (see **Table 4**). Its eigenvector centrality is 0.7397 which, although not the maximum, is very close to it (see **Table 4**). Indeed, MRE11 is an integral part of the protein complex of RAD50-MRE11A-NBS1 known as the MRN complex (Porras, 2014; Shibata et al., 2014; Mukherjee et al., 2019). It plays a key role in homologous recombination, and it is generally believed that MRE11 initiates double-strand breaks resection. In particular, the authors show that the loss of MRE11 reduces the efficiency of homologous recombination in human TK6 cells without affecting double-strand breaks resection, indicating a role for MRE11 in homologous recombination also at a post-resection step.

The high value of the eigenvector centrality fork BLM, RPA2 and MRE11 confirms the crucial role of these proteins in the network and expresses the fact that they are pointed by

FIGURE 10 | Mean of the sensitivity index distributions for the proteins in HR network Orlic-Milacic, (2015) merged with FANCM pathway (Pathways Commons, 2022). These results refer to simulation in the time interval [0, 1400] a.u., and initial values of the proteins randomly sampled in the range [18, 20] a. u. and kinetics rate values sample in [$10^{-5}$, $10^{-6}$] a. u.

nodes that have a high value of the eigenvector centrality too. Indeed, if a node is pointed to by many nodes (which also have high eigenvector centrality) then that node will have high eigenvector centrality (Fletcher and Wennekers, 2018). The high sensitivity to the parameters characterising the dynamics of the interactions between these proteins and the partners pointing to them indicates the great influence that these partner nodes have on these proteins. Interestingly, RAD51 does not result sensitive to any parameter. The RAD51 encodes a protein that is essential for repairing damaged DNA. Recent findings have indicated RAD51 protein is overexpressed in a variety of tumours Chen et al. (2017). The overexpression of RAD51 causes improper and hyper-recombination, and thus contributes to genomic instability and



FIGURE 11 | Coefficient of variation of the sensitivity index distributions for the proteins in HR network (Orlic-Milacic, 2015) merged with FANCM pathway (Pathways Commons, 2022). These results refer to simulation in the time interval [18, 20] a.u., and initial values of the proteins randomly sampled in the range [$10^{-5}$, $10^{-6}$] a. u.

genetic diversity. Genomic instability might, in turn, drive regular cells towards neoplastic transformation or further contributes to cancer metastatic progression (Chen et al., 2007). The RAD51 protein binds to the DNA at the site of a break and encapsulates it in a protein sheath, initiating the repair process MedlinePlus (2022); Uniprot (2022). RAD51 protein interacts with BRCA1 and BRCA2, to fix damaged DNA. The BRCA2 protein regulates the activity of the RAD51 protein by transporting it in the nucleus to sites of DNA damage. Although the interaction between the BRCA1 protein and the RAD51 protein has still to be elucidate, research suggests that BRCA1 may also activate RAD51 in response to DNA damage (Cousineau et al., 2005; Chappell et al., 2016). The result of the sensitivity analysis found seems to contradict the important role of this protein in these interactions. Indeed, for example, one might expect a high sensitivity of RAD51 to the k139 due to its interaction with BRCA2 (see **Supplementary Table S2**). One explanation for this contradiction could be that since the mechanisms of interaction of RAD51 with these proteins are not fully known, the model used in this study could be an oversimplification of the interaction of RAD51 with its partners. If more accurate models in the future confirm the low sensitivity of RAD51 to the parameters of the rate equations describing the dynamics of the network, it will be necessary to investigate the physical and biological characteristics that make it so stable to perturbations. The fact that RAD51 has a low value of vibrational centrality in this study is a factor in favour of the possible confirmation of this case.

In **Figures 7**, **8** we report the results of the sensitivity analysis obtained selecting different ranges of initial conditions and parameters values. The plots highlights RAD51C, MRE11, RAD50 and BRCA2 as the most sensitive nodes to the parameters. We comment in the Section Remarks the expected differences and similarities in the results of sensitivity analysis when we change the intervals of initial conditions and parameters.

## 4.1 Analysis of the HR Pathway Merged With FANCM Pathway

We repeated the analysis on the HR's network extended by adding the pathways of FANCM gene (Fanconi Anaemia Group M Protein), obtained from Pathways Commons (Pathways Commons, 2022) given its important role in genome duplication, repair mechanisms and its involvement in the development of Fanconi anaemia, which several studies report to be a syndrome related to cancer predisposition (Deans and West, 2009; Xue et al., 2015; Bhattacharjee and Nandi, 2017; Pan et al., 2017; Wang et al., 2018). Finally, a recent study of Panday et al. reports that FANCM regulates repair pathway choice at stalled replication forks (Ling et al., 2016; Panday et al., 2021). FANCM and BLM have a similar role and cooperatively act in the DNA repair mechanisms (Panday et al., 2021), and through this analysis we want to investigate on this similarity.

The new network including HR and FANCM pathways is made up of 46 nodes and 316 edges. The left part of **Figure 4**

shows the FANCM pathway and its connection with HR network. As reported in **Table 8**, the calculation of centrality measures led to the following results:

- BLM has the highest score in vibrational centrality
- FANCM has the highest score in total degree, in degree and betweenness
- RAD50 scores first for subgraph centrality
- RAD52 scores first for clustering coefficient
- RPA1 scores first for out-degree
- RPA2 scores first for eigenvector centrality
- RPA3 scores first for hub centrality
- RPA4 scores first for clustering coefficient
- SEM1 scores first for clustering coefficient.

The distribution of the centrality measures on the entire network is shown in **Figure 9**. These results not only re-emphasise as central the genes/proteins already identified in the analysis of the HR network alone in (Orlic-Milacic, 2015), but also highlight the central role of FANCM and as a node of particular relevance due to their high in-degree and high betweenness. By assigning an initial quantity between 18 and 20 (expressed in arbitrary units) to each node, the solution of the system of 46 differential equations converges for values of rate constants in the range $10^{-4}$ and $10^{-5}$ a.u. In file Simulations_of_Dynamics_HR_FANCM_pathway.pdf provided in the Supplementary Material, we show the time evolution curves (obtained as a solution of the equations) of each node of the HR pathway merged with FANCM pathway. The parameter sensitivity analysis was conducted by perturbing each parameter in the convergence range of the solution and yielded the results shown in **Figures 10**, **11**. We found that the nodes most sensitive to the parameters are RAD50, NBN, BRCA2, MRE11 and RAD51B. Compared to what was obtained in the analysis of the HR network alone, we find here that BLM is no longer at the top in terms of parameter sensitivity while still maintaining a central role in terms of vibration centrality.

## 4.2 Remarks

The method provides a range of values for the rate constants within which the solution to the problem exists at the given set of initial values for the node concentration/abundance. We can interpret this range as that of the 'most probable' range of values if.

- the initial conditions are known
- the analysed network does not exclude important interactions occurring *in vivo* and if the system is subjected to the conditions of the real system *in vitro* and *in vivo*. The network considered is only an extract of a much more complex network (still not completely known) that operates *in vivo* and in interaction with environmental factors.

We also note that we do not dispose of experimental time curves that can be used to calibrate the model. Calibration from

experimental data rather than sensitivity analysis would be the most appropriate method to use to obtain an estimated (even interval) parameter estimate. In the absence of both experimental data. We agree with the Reviewer that sensitivity analysis provides information on the minimum set of parameters to be inferred from experimental data, since the parameters to which the model is less sensitive are less influential.

Finally, we also observe that having fixed a set of initial values for the concentrations/abundances of the network components, more than one set of intervals for the rate constants could guarantee the convergence of the numerical method of solving the system of differential equations. It is also true that by changing the initial values of the concentrations, the range of values of the rate constants for which the system converges could change. The results that we report in this new version of the manuscript show, for example, that if the range of the initial concentration values is between 0 and 100, the numerical solution is found for rate constant values between 0 and 0.01, whereas if the range of the initial concentration values is a few tens, the numerical solution is found for rate constant values between $10^{-6}$ and 106–5. A reduction of 10 in the order of magnitude of the initial concentration values thus corresponds to a reduction of $10^{-3}$ in the order of magnitude of the rate constants. This is an indication that the system is underdetermined, and in fact consists of more parameters than the number of variables and in the complete absence of experimental data. All this also shows that calibrating the model in the light of experimental data is the best way to hope for a set of 'probable' values of the rate constants.

The work shown in this study therefore does not so much emphasise the numerical solutions, but, through a mathematical model, wants to test the susceptibility of the network components to the parameters and wants to integrate it with the role that the network components have (estimated by the centrality measurements).

# 5 CONCLUSION

This report presented an application of network analysis and mathematical modelling to the double-strand break repair pathway homologous recombination repair (HR). The complexity of the network of repair mechanisms itself, as well as the complexity of its interactions with the surrounding environment (Li et al., 2009; Chatterjee and Walker, 2017; Kusakabe et al., 2019; Poetsch, 2020; Roux et al., 2021), and the mutations of its components make its mathematical modelling particularly difficult, especially when based on rate equations. It is therefore of great necessity to have a tool that can implement these two important steps:

1. network analysis including standard centrality measures and new measures to quantify the robustness and responsiveness of the network to stimuli and stresses not dependent on the network topology
2. automatic construction of a mathematical model, for its analysis, and which allows to carry out refinements and

modifications, when new data and new experimental knowledge make it necessary.

The implementation of these step is an innovative perspective for the analysis of DNA repair mechanisms. So far in the literature, there are many studies and analyses focused on the genetic and genomic aspects of the pathways, but studies on the mathematical modelling of its dynamics are absent. Our study therefore aims to fill this gap, since the mechanisms of DNA repair are governed by genes, proteins and pathways in continuous communication with the environment. For this reason, the analysis of the dynamics of the network is particularly useful, since it can quantify the vulnerability of the network and the modes of response to stimuli and exogenous stress. To the best of our knowledge there are no schemes for translating a graph associated with a biological network into a set of dynamic equations. The main reason for this is that there is no unambiguously defined semantics of a graphic representation of a biological network, i.e. there is no unambiguous definition of the graphic symbolism in terms of the mathematical equation describing the interaction indicated by that symbolism. The translation model we propose in this study is a basic model that describes the interactions indicated by the graph with linear first-order differential equations in the parameters. The code that implements this translation, however, gives the user the possibility to modify the model where he/she deems it appropriate in the light of available biological knowledge, or in the case he/she like to generate new hypothetical scenarios. We believe that the availability of a tool such as NADS can support the investigation of such a complex network that is subject to continuous interaction with external agents, not only to understand its dynamics, but also to predict its evolution and identify points of vulnerability for the benefit of the medical applications that this research may provide.

# DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors. the software used int eh analysis is available at https://gitlab.inf.unibz.it/Paola.Lecca/network-analyzer-and-dynamics-simulator.

# AUTHOR CONTRIBUTIONS

PL designed the methods and implemented the R codes for the analysis of the biological network, the automatic translation of the hypergraph into a set of rate equations, the numerical solution of the model and the sensitivity analysis. AI-N provided the biological background and provided the interpretation of the results obtained, in the light of an extensive literature search. Both authors contributed to the conceptualisation of the paper, its writing and revision.

# FUNDING

# SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.878148/full#supplementary-material

# REFERENCES

Alberts, B. (2015). *Molecular biology of the Cell.* New York, NY: Garland Science, Taylor and Francis Group.

Ashtiani, M., Salehzadeh-Yazdi, A., Razaghi-Moghadam, Z., Hennig, H., Wolkenhauer, O., Mirzaie, M., et al. (2018). A Systematic Survey of Centrality Measures for Protein-Protein Interaction Networks. *BMC Syst. Biol.* 12, 80. doi:10.1186/s12918-018-0598-2

Bass, T. E., Luzwick, J. W., Kavanaugh, G., Carroll, C., Dungrawala, H., Glick, G. G., et al. (2016). ETAA1 Acts at Stalled Replication Forks to Maintain Genome Integrity. *Nat. Cell Biol.* 18, 1185–1195. doi:10.1038/ncb3415

Beikzadeh, M., and Latham, M. P. (2021). The Dynamic Nature of the Mre11-Rad50 DNA Break Repair Complex. *Prog. Biophysics Mol. Biol.* 163, 14–22. doi:10.1016/j.pbiomolbio.2020.10.007

Bennett, C. B., Lewis, A. L., Baldwin, K. K., and Resnick, M. A. (1993). Lethality Induced by a Single Site-specific Double-Strand Break in a Dispensable Yeast Plasmid. *Proc. Natl. Acad. Sci. U.S.A.* 90, 5613–5617. doi:10.1073/pnas.90.12.5613

Bhattacharjee, S., and Nandi, S. (2017). DNA Damage Response and Cancer Therapeutics through the Lens of the Fanconi Anemia DNA Repair Pathway. *Cell Commun. Signal* 15, 41. doi:10.1186/s12964-017-0195-9

Bian, L., Meng, Y., Zhang, M., and Li, D. (2019). MRE11-RAD50-NBS1 Complex Alterations and DNA Damage Response: Implications for Cancer Treatment. *Mol. Cancer* 18, 169. doi:10.1186/s12943-019-1100-5

Bouwman, P., and Jonkers, J. (2012). The Effects of Deregulated DNA Damage Signalling on Cancer Chemotherapy Response and Resistance. *Nat. Rev. Cancer* 12, 587–598. doi:10.1038/nrc3342

Branzei, D., and Foiani, M. (2008). Regulation of DNA Repair throughout the Cell Cycle. *Nat. Rev. Mol. Cell Biol.* 9, 297–308. doi:10.1038/nrm2351

Brosey, C. A., Yan, C., Tsutakawa, S. E., Heller, W. T., Rambo, R. P., and Tainer, J. A. (2013). A new structural framework for integrating replication protein A into DNA processing machinery. *Nucleic Acids Res.* 41 (4), 2313–2327. doi:10.1093/nar/gks1332

Bugreev, D. V., Mazina, O. M., and Mazin, A. V. (2006). Rad54 Protein Promotes Branch Migration of Holliday Junctions. *Nature* 442, 590–593. doi:10.1038/nature04889

Çağlayan, M., and Wilson, S. H. (2015). Oxidant and Environmental Toxicant-Induced Effects Compromise DNA Ligation during Base Excision DNA Repair. *DNA Repair* 35, 85–89. doi:10.1016/j.dnarep.2015.09.010

Carney, J. P., Maser, R. S., Olivares, H., Davis, E. M., Le Beau, M., Yates, J. R., 3rd, et al. (1998). The Hmre11/hrad50 Protein Complex and Nijmegen Breakage Syndrome: Linkage of Double-Strand Break Repair to the Cellular DNA Damage Response. *Cell* 93, 477–486. doi:10.1016/s0092-8674(00)81175-7

Carusillo, A., and Mussolino, C. (2020). DNA Damage: From Threat to Treatment. *Cells* 9, 1665. doi:10.3390/cells9071665

Cerami, E. G., Gross, B. E., Demir, E., Rodchenkov, I., Babur, O., Anwar, N., et al. (2010). Pathway Commons, a Web Resource for Biological Pathway Data. *Nucleic Acids Res.* 39, D685–D690. doi:10.1093/nar/gkq1039

Cerbinskaite, A., Mukhopadhyay, A., Plummer, E. R., Curtin, N. J., and Edmondson, R. J. (2012). Defective Homologous Recombination in Human Cancers. *Cancer Treat. Rev.* 38, 89–100. doi:10.1016/j.ctrv.2011.04.015

Chappell, W. H., Gautam, D., Ok, S. T., Johnson, B. A., Anacker, D. C., and Moody, C. A. (2016). Homologous Recombination Repair Factors Rad51 and BRCA1 Are Necessary for Productive Replication of Human Papillomavirus 31. *J. Virol.* 90, 2639–2652. doi:10.1128/jvi.02495-15

Chatterjee, N., and Walker, G. C. (2017). Mechanisms of DNA Damage, Repair, and Mutagenesis. *Environ. Mol. Mutagen.* 58, 235–263. doi:10.1002/em.22087

Chen, C.-Y., Ping, Y.-H., Lee, H.-C., Chen, K.-H., Lee, Y.-M., Chan, Y.-J., et al. (2007). Open Reading Frame 8a of the Human Severe Acute Respiratory Syndrome Coronavirus Not Only Promotes Viral Replication but Also Induces Apoptosis. *J. Infect. Dis.* 196, 405–415. doi:10.1086/519166

Chen, Q., Cai, D., Li, M., and Wu, X. (2017). The Homologous Recombination Protein RAD51 Is a Promising Therapeutic Target for Cervical Carcinoma. *Oncol. Rep.* 38, 767–774. doi:10.3892/or.2017.5724

Chu, W. K., and Hickson, I. D. (2009). RecQ Helicases: Multifunctional Genome Caretakers. *Nat. Rev. Cancer* 9, 644–654. doi:10.1038/nrc2682

Cooper, G. (2000). *The Cell : A Molecular Approach.* Washington, D.C. Sunderland, Mass: ASM Press Sinauer Associates.

Cousineau, I., Abaji, C., and Belmaaza, A. (2005). BRCA1 Regulates RAD51 Function in Response to DNA Damage and Suppresses Spontaneous Sister Chromatid Replication Slippage: Implications for Sister Chromatid Cohesion, Genome Stability, and Carcinogenesis. *Cancer Res.* 65, 11384–11391. doi:10.1158/0008-5472.can-05-2156

De Bont, R. (2004). Endogenous DNA Damage in Humans: a Review of Quantitative Data. *Mutagenesis* 19, 169–185. doi:10.1093/mutage/geh025

de Jager, M., van Noort, J., van Gent, D. C., Dekker, C., Kanaar, R., and Wyman, C. (2001). Human Rad50/Mre11 Is a Flexible Complex that Can Tether DNA Ends. *Mol. Cell* 8, 1129–1135. doi:10.1016/s1097-2765(01)00381-1

Deans, A. J., and West, S. C. (2009). FANCM Connects the Genome Instability Disorders Bloom's Syndrome and Fanconi Anemia. *Mol. Cell* 36, 943–953. doi:10.1016/j.molcel.2009.12.006

Dietlein, F., and Reinhardt, H. C. (2014). Molecular Pathways: Exploiting Tumor-specific Molecular Defects in DNA Repair Pathways for Precision Cancer Therapy. *Clin. Cancer Res.* 20, 5882–5887. doi:10.1158/1078-0432.ccr-14-1165

Estrada, E., and Hatano, N. (2010). A Vibrational Approach to Node Centrality and Vulnerability in Complex Networks. *Phys. A Stat. Mech. its Appl.* 389, 3648–3660. doi:10.1016/j.physa.2010.03.030

Estrada, E., and Rodríguez-Velázquez, J. A. (2005). Subgraph Centrality in Complex Networks. *Phys. Rev. E* 71, 056103. doi:10.1103/PhysRevE.71.056103

Estrada, E., and Ross, G. J. (2018). Centralities in Simplicial Complexes. Applications to Protein Interaction Networks. *J. Theor. Biol.* 438, 46–60. doi:10.1016/j.jtbi.2017.11.003

Estrada, E. (2011). *The Structure of Complex Networks.* New York, United States: Oxford University Press.

Fanning, E. (2006). A Dynamic Model for Replication Protein a (RPA) Function in DNA Processing Pathways. *Nucleic Acids Res.* 34, 4126–4137. doi:10.1093/nar/gkl550

Feldkamp, M. D., Mason, A. C., Eichman, B. F., and Chazin, W. J. (2014). Structural Analysis of Replication Protein a Recruitment of the DNA Damage Response Protein SMARCAL1. *Biochemistry* 53, 3052–3061. doi:10.1021/bi500252w

Fletcher, J. M., and Wennekers, T. (2018). From Structure to Activity: Using Centrality Measures to Predict Neuronal Activity. *Int. J. Neur. Syst.* 28, 1750013. doi:10.1142/s0129065717500137

Foertsch, F., Kache, T., Drube, S., Biskup, C., Nasheuer, H. P., and Melle, C. (2019). Determination of the Number of RAD51 Molecules in Different Human Cell Lines. *Cell Cycle* 18, 3581–3588. doi:10.1080/15384101.2019.1691802

Fornito, A., Zalesky, A., and Bullmore, E. T. (2016). "Centrality and Hubs," in *Fundamentals of Brain Network Analysis* (San Diego: Academic Press), 137–161. doi:10.1016/B978-0-12-407908-3.00005-4

Fourtziala, E., Givalos, N., Alexakis, N., Griniatsos, J., Alevizopoulos, N., Kavantzas, N., et al. (2020). Replication Protein A (RPA1, RPA2 and RPA3) Expression in Gastric Cancer: Correlation with Clinicopathologic Parameters and Patients' Survival. *J. BUON* 25, 1482–1489.

Friedberg, E. C., Aguilera, A., Gellert, M., Hanawalt, P. C., Hays, J. B., Lehmann, A. R., et al. (2006). DNA Repair: From Molecular Mechanism to Human Disease. *DNA Repair* 5, 986–996. doi:10.1016/j.dnarep.2006.05.005

Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., and Ellenberger, T. (2005). *DNA Repair and Mutagenesis*. Washington, DC: ASM Press.

Ganai, R. A., and Johansson, E. (2016). DNA Replication-A Matter of Fidelity. *Mol. Cell* 62, 745–755. doi:10.1016/j.molcel.2016.05.003

[Dataset] GeneCard (2022). *GeneCards – the Human Gene Database*. Available at: https://www.genecards.org/.

Ghasemi, M., Seidkhani, H., Tamimi, F., Rahgozar, M., and Masoudi-Nejad, A. (2014). Centrality Measures in Biological Networks. *Curr. Bioinforma.* 9, 426–441. doi:10.2174/15748936113086660013

Ghosal, G., and Chen, J. (2013). DNA Damage Tolerance: a Double-Edged Sword Guarding the Genome. *Transl. Cancer Res.* 2, 107–129. doi:10.3978/j.issn.2218-676X.2013.04.01

Gravel, S., Chapman, J. R., Magill, C., and Jackson, S. P. (2008). DNA Helicases Sgs1 and BLM Promote DNA Double-Strand Break Resection. *Genes Dev.* 22, 2767–2772. doi:10.1101/gad.503108

Haahr, P., Hoffmann, S., Tollenaere, M. A. X., Ho, T., Toledo, L. I., Mann, M., et al. (2016). Activation of the ATR Kinase by the RPA-Binding Protein ETAA1. *Nat. Cell Biol.* 18, 1196–1207. doi:10.1038/ncb3422

Hansen, D. L., Shneiderman, B., Smith, M. A., and Himelboim, I. (2020). "Calculating and Visualizing Network Metrics," in *Analyzing Social Media Networks with NodeXL* (Elsevier), 79–94. doi:10.1016/b978-0-12-817756-3.00006-6

Hoeijmakers, J. H. J. (2001). DNA Repair Mechanisms. *Maturitas* 38, 17–22. doi:10.1016/s0378-5122(00)00188-2

Hoeijmakers, J. H. J. (2009). DNA Damage, Aging, and Cancer. *N. Engl. J. Med.* 361, 1475–1485. doi:10.1056/nejmra0804615

Hosoya, N., and Miyagawa, K. (2014). Targeting DNA Damage Response in Cancer Therapy. *Cancer Sci.* 105, 370–388. doi:10.1111/cas.12366

Huang, R., and Zhou, P.-K. (2021). DNA Damage Repair: Historical Perspectives, Mechanistic Pathways and Clinical Translation for Targeted Cancer Therapy. *Sig Transduct. Target Ther.* 6, 254. doi:10.1038/s41392-021-00648-7

Huertas, P. (2010). DNA Resection in Eukaryotes: Deciding How to Fix the Break. *Nat. Struct. Mol. Biol.* 17, 11–16. doi:10.1038/nsmb.1710

Huertas, P., and Jackson, S. P. (2009). Human CtIP Mediates Cell Cycle Control of DNA End Resection and Double Strand Break Repair. *J. Biol. Chem.* 284, 9558–9565. doi:10.1074/jbc.m808906200

[Dataset] Human Protein Atlas (2022). *Human Protein Atlas*.

Jackson, S. P., and Bartek, J. (2009). The DNA-Damage Response in Human Biology and Disease. *Nature* 461, 1071–1078. doi:10.1038/nature08467

Jalili, M., Salehzadeh-Yazdi, A., Gupta, S., Wolkenhauer, O., Yaghmaie, M., Resendis-Antonio, O., et al. (2016). Evolution of Centrality Measurements for the Detection of Essential Proteins in Biological Networks. *Front. Physiol.* 7, 375. doi:10.3389/fphys.2016.00375

Kaur, E., Agrawal, R., and Sengupta, S. (2021). Functions of BLM Helicase in Cells: Is it Acting like a Double-Edged Sword? *Front. Genet.* 12, 634789. doi:10.3389/fgene.2021.634789

Kawale, A. S., and Sung, P. (2020). Mechanism and Significance of Chromosome Damage Repair by Homologous Recombination. *Essays Biochem.* 64, 779–790. doi:10.1042/ebc20190093

Kennedy, R. D., and D'Andrea, A. D. (2006). DNA Repair Pathways in Clinical Practice: Lessons from Pediatric Cancer Susceptibility Syndromes. *J. Clin. Oncol.* 24, 3799–3808. doi:10.1200/jco.2005.05.4171

Keshav, K. F., Chen, C., and Dutta, A. (1995). Rpa4, a Homolog of the 34-kilodalton Subunit of the Replication Protein a Complex. *Mol. Cell Biol.* 15, 3119–3128. doi:10.1128/mcb.15.6.3119

Kim, C., Paulus, B. F., and Wold, M. S. (1994). Interactions of Human Replication Protein a with Oligonucleotides. *Biochemistry* 33, 14197–14206. doi:10.1021/bi00251a031

Koschützki, D., and Schreiber, F. (2008). Centrality Analysis Methods for Biological Networks and Their Application to Gene Regulatory Networks. *Gene Regul. Syst. Bio* 2, GRSB.S702. GRSB. doi:10.4137/grsb.s702

Kovalchuk, I. (2016). *Genome Stability*. San Diego, CA: Elsevier, 1–18.

Kowalczykowski, S. C. (2015). An Overview of the Molecular Mechanisms of Recombinational DNA Repair. *Cold Spring Harb. Perspect. Biol.* 7, a016410. doi:10.1101/cshperspect.a016410

Krogan, N. J., Lam, M. H. Y., Fillingham, J., Keogh, M.-C., Gebbia, M., Li, J., et al. (2004). Proteasome Involvement in the Repair of DNA Double-Strand Breaks. *Mol. Cell* 16, 1027–1034. doi:10.1016/j.molcel.2004.11.033

Kusakabe, M., Onishi, Y., Tada, H., Kurihara, F., Kusao, K., Furukawa, M., et al. (2019). Mechanism and Regulation of DNA Damage Recognition in Nucleotide Excision Repair. *Genes Environ* 41, 2. doi:10.1186/s41021-019-0119-6

Le, H. P., Ma, X., Vaquero, J., Brinkmeyer, M., Guo, F., Heyer, W.-D., et al. (2020). DSS1 and ssDNA Regulate Oligomerization of BRCA2. *Nucleic Acids Res.* 48, 7818–7833. doi:10.1093/nar/gkaa555

Lecca, P., Mura, I., Re, A., Barker, G. C., and Ihekwaba, A. E. (2016). Time Series Analysis of the bacillus Subtilis Sporulation Network Reveals Low Dimensional Chaotic Dynamics. *Front. Microbiol.* 7, 1760. doi:10.3389/fmicb.2016.01760

Lecca, P., and Re, A. (2019). *Theoretical Physics for Biological Systems*. London, England: CRC Press.

Li, L., di Guan, Y., sha Chen, X., ming Yang, J., and Cheng, Y. (2021). DNA Repair Pathways in Cancer Therapy and Resistance. *Front. Pharmacol.* 11, 629266. doi:10.3389/fphar.2020.629266

Li, Y., Marion, M.-J., Zipprich, J., Santella, R. M., Freyer, G., and Brandt-Rauf, P. W. (2009). Gene-Environment Interactions Between DNA Repair Polymorphisms and Exposure to the Carcinogen Vinyl Chloride. *Biomarkers* 14, 148–155. doi:10.1080/13547500902811266

Lin, Y.-L., Shivji, M. K. K., Chen, C., Kolodner, R., Wood, R. D., and Dutta, A. (1998). The Evolutionarily Conserved Zinc Finger Motif in the Largest Subunit of Human Replication Protein a Is Required for DNA Replication and Mismatch Repair but Not for Nucleotide Excision Repair. *J. Biol. Chem.* 273, 1453–1461. doi:10.1074/jbc.273.3.1453

Lindahl, T. (1993). Instability and Decay of the Primary Structure of DNA. *Nature* 362, 709–715. doi:10.1038/362709a0

Ling, C., Huang, J., Yan, Z., Li, Y., Ohzeki, M., Ishiai, M., et al. (2016). Bloom Syndrome Complex Promotes FANCM Recruitment to Stalled Replication Forks and Facilitates Both Repair and Traverse of DNA Interstrand Crosslinks. *Cell Discov.* 2, 16047. doi:10.1038/celldisc.2016.47

Liu, J., Doty, T., Gibson, B., and Heyer, W.-D. (2010). Human BRCA2 Protein Promotes RAD51 Filament Formation on RPA-Covered Single-Stranded DNA. *Nat. Struct. Mol. Biol.* 17, 1260–1262. doi:10.1038/nsmb.1904

Liu, T., and Huang, J. (2016). DNA End Resection: Facts and Mechanisms. *Genomics, Proteomics Bioinforma.* 14, 126–130. doi:10.1016/j.gpb.2016.05.002

Liu, Y. (2001). p53 Protein at the Hub of Cellular DNA Damage Response Pathways through Sequence-specific and Non-sequence-specific DNA Binding. *Carcinogenesis* 22, 851–860. doi:10.1093/carcin/22.6.851

Lok, B. H., and Powell, S. N. (2012). Molecular Pathways: Understanding the Role of Rad52 in Homologous Recombination for Therapeutic Advancement. *Clin. Cancer Res.* 18, 6400–6406. doi:10.1158/1078-0432.ccr-11-3150

Lu, H., and Davis, A. J. (2021). Human RecQ Helicases in DNA Double-Strand Break Repair. *Front. Cell Dev. Biol.* 9, 640755. doi:10.3389/fcell.2021.640755

Maréchal, A., and Zou, L. (2014). RPA-Coated Single-Stranded DNA as a Platform for Post-translational Modifications in the DNA Damage Response. *Cell Res.* 25, 9–23. doi:10.1038/cr.2014.147

Marsden, P. V. (2005). "Network Analysis," in *Encyclopedia of Social Measurement* (Elsevier), 819–825. doi:10.1016/b0-12-369398-5/00409-6

[Dataset] MedlinePlus (2022). MedlinePlus Data Base. Available at: https://medlineplus.gov/genetics/gene/rad51/(Accessed 02 10.2022)

Mukherjee, S., Abdisalaam, S., Bhattacharya, S., Srinivasan, K., Sinha, D., and Asaithamby, A. (2019). "Mechanistic Link between DNA Damage Sensing, Repairing and Signaling Factors and Immune Signaling," in *DNA Repair* (Elsevier), 297–324. doi:10.1016/bs.apcsb.2018.11.004

[Dataset] National library of Medicine (2022). *National Library of Medicine*. National Center for Biotechnology Information. Available at: https://www.nlm.nih.gov/.

Nguyen, G. H., Dexheimer, T. S., Rosenthal, A. S., Chu, W. K., Singh, D. K., Mosedale, G., et al. (2013). A Small Molecule Inhibitor of the BLM Helicase Modulates Chromosome Stability in Human Cells. *Chem. Biol.* 20, 55–62. doi:10.1016/j.chembiol.2012.10.016

Nimonkar, A. V., Özsoy, A. Z., Genschel, J., Modrich, P., and Kowalczykowski, S. C. (2008). Human Exonuclease 1 and BLM Helicase Interact to Resect DNA and Initiate DNA Repair. *Proc. Natl. Acad. Sci. U.S.A.* 105, 16906–16911. doi:10.1073/pnas.0809380105

[Dataset] Orlic-Milacic, M. (2015). HDR through Homologous Recombination (Hrr). Available at: https://apps.pathwaycommons.org/search?q=homologous%20recombinations&type=Pathway (Accessed 10 01.2021)

Pan, X., Drosopoulos, W. C., Sethi, L., Madireddy, A., Schildkraut, C. L., and Zhang, D. (2017). FANCM, BRCA1, and BLM Cooperatively Resolve the Replication Stress at the ALT Telomeres. *Proc. Natl. Acad. Sci. U.S.A.* 114, E5940. doi:10.1073/pnas.1708065114

Panday, A., Willis, N. A., Elango, R., Menghi, F., Duffey, E. E., Liu, E. T., et al. (2021). FANCM Regulates Repair Pathway Choice at Stalled Replication Forks. *Mol. Cell* 81, 2428–2444. e6. doi:10.1016/j.molcel.2021.03.044

[Dataset] Pathways Commons (2022). Pathways Commons. .Available at: https://apps.pathwaycommons.org/search?type=Pathway&q=FANCM

Poetsch, A. R. (2020). The Genomics of Oxidative DNA Damage, Repair, and Resulting Mutagenesis. *Comput. Struct. Biotechnol. J.* 18, 207–219. doi:10.1016/j.csbj.2019.12.013

Porras, O. (2014). "DNA Repair Defects," in *Stiehm's Immune Deficiencies*. Editors K. E. Sullivan and E. R. Stiehm (Amsterdam: Academic Press), 199–219. doi:10.1016/b978-0-12-405546-9.00007-8

[Dataset] R Studio (2022). R Studio. Available at: https://rstudio.com/.

[Dataset] Reed, S. H., and Waters, R. (2005). DNA Repair. *eLS*. doi:10.1038/npg.els.0005284

Ross, C. A., and Truant, R. (2016). A Unifying Mechanism in Neurodegeneration. *Nature* 541, 34–35. doi:10.1038/nature21107

Rossi, M. J., DiDomenico, S. F., Patel, M., and Mazin, A. V. (2021). RAD52: Paradigm of Synthetic Lethality and New Developments. *Front. Genet.* 12, 780293. doi:10.3389/fgene.2021.780293

Roux, P., Salort, D., and Xu, Z. (2021). Adaptation to DNA Damage as a Bet-Hedging Mechanism in a Fluctuating Environment. *R. Soc. open sci.* 8, 210460. doi:10.1098/rsos.210460

Safran, M., Rosen, N., Twik, M., BarShir, R., Stein, T. I., Dahary, D., et al. (2021). "The GeneCards Suite," in *Practical Guide to Life Science Databases* (Springer Singapore), 27. doi:10.1007/978-981-16-5812-9_2

Sancar, A., Lindsey-Boltz, L. A., Ünsal-Kaçmaz, K., and Linn, S. (2004). Molecular Mechanisms of Mammalian DNA Repair and the DNA Damage Checkpoints. *Annu. Rev. Biochem.* 73, 39–85. doi:10.1146/annurev.biochem.73.011303.073723

Shibata, A., Moiani, D., Arvai, A. S., Perry, J., Harding, S. M., Genois, M.-M., et al. (2014). DNA Double-Strand Break Repair Pathway Choice Is Directed by Distinct MRE11 Nuclease Activities. *Mol. Cell* 53, 7–18. doi:10.1016/j.molcel.2013.11.003

Sone, T., Saeki, Y., Toh-e, A., and Yokosawa, H. (2004). Sem1p Is a Novel Subunit of the 26 S Proteasome from saccharomyces Cerevisiae. *J. Biol. Chem.* 279, 28807–28816. doi:10.1074/jbc.m403165200

Stefanovie, B., Hengel, S. R., Mlcouskova, J., Prochazkova, J., Spirek, M., Nikulenkov, F., et al. (2019). DSS1 Interacts with and Stimulates RAD52 to Promote the Repair of DSBs. *Nucleic Acids Res.* 48, 694–708. doi:10.1093/nar/gkz1052

Sugiyama, T., and Kantake, N. (2009). Dynamic Regulatory Interactions of Rad51, Rad52, and Replication Protein-A in Recombination Intermediates. *J. Mol. Biol.* 390, 45–55. doi:10.1016/j.jmb.2009.05.009

Sun, Y., McCorvie, T. J., Yates, L. A., and Zhang, X. (2019). Structural Basis of Homologous Recombination. *Cell. Mol. Life Sci.* 77, 3–18. doi:10.1007/s00018-019-03365-1

[Dataset] Uniprot (2022). Uniprot Database. Available at: https://www.uniprot.org/uniprot/Q06609 (Accessed 02 10.2022)

van Brabant, A. J., Ye, T., Sanz, M., German, J. L., Ellis, N. A., and Holloman, W. K. (2000). Binding and Melting of D-Loops by the Bloom Syndrome Helicase. *Biochemistry* 39, 14617–14625. doi:10.1021/bi0018640

Wang, H., Li, S., Oaks, J., Ren, J., Li, L., and Wu, X. (2018). The Concerted Roles of FANCM and Rad52 in the Protection of Common Fragile Sites. *Nat. Commun.* 9, 2791. doi:10.1038/s41467-018-05066-y

Wang, P., Lü, J., and Yu, X. (2014). Identification of Important Nodes in Directed Biological Networks: A Network Motif Approach. *Plos One* 9, e106132. doi:10.1371/journal.pone.0106132

Wolters, S., and Schumacher, B. (2013). Genome Maintenance and Transcription Integrity in Aging and Disease. *Front. Genet.* 4, 19. doi:10.3389/fgene.2013.00019

Wood, R. D. (2018). Fifty Years since DNA Repair Was Linked to Cancer. *Nature* 557, 648–649. doi:10.1038/d41586-018-05255-1

Wood, R. D., Mitchell, M., and Lindahl, T. (2005). Human DNA Repair Genes, 2005. *Mutat. Research/Fundamental Mol. Mech. Mutagen.* 577, 275–283. doi:10.1016/j.mrfmmm.2005.03.007

Xue, X., Sung, P., and Zhao, X. (2015). Functions and Regulation of the Multitasking FANCM Family of DNA Motor Proteins. *Genes Dev.* 29, 1777–1788. doi:10.1101/gad.266593.115

Yadav, S., Anbalagan, M., Baddoo, M., Chellamuthu, V. K., Mukhopadhyay, S., Woods, C., et al. (2020). Somatic Mutations in the DNA Repairome in Prostate Cancers in African Americans and Caucasians. *Oncogene* 39, 4299–4311. doi:10.1038/s41388-020-1280-x

Yang, Y., Dou, S.-X, Xu, Y.-N., Bazeille, N., Wang, P.-Y., Rigolet, P., et al. (2010). Kinetic Mechanism of DNA Unwinding by the BLM Helicase Core and Molecular Basis for its Low Processivity. *Biochemistry* 49, 656–668. doi:10.1021/bi901459c

Yousefzadeh, M., Henpita, C., Vyas, R., Soto-Palma, C., Robbins, P., and Niedernhofer, L. (2021). DNA Damage-How and Why We Age? *eLife* 10, e62852. doi:10.7554/eLife.62852

Yun, M. H., and Hiom, K. (2009). CtIP-BRCA1 Modulates the Choice of DNA Double-Strand-Break Repair Pathway throughout the Cell Cycle. *Nature* 459, 460–463. doi:10.1038/nature07955

Zhao, F., Kim, W., Kloeber, J. A., and Lou, Z. (2020). DNA End Resection and its Role in DNA Replication and DSB Repair Choice in Mammalian Cells. *Exp. Mol. Med.* 52, 1705–1714. doi:10.1038/s12276-020-00519-1

Zhao, W., Vaithiyalingam, S., San Filippo, J., Maranon, D. G., Jimenez-Sainz, J., Fontenay, G. V., et al. (2015). Promotion of BRCA2-dependent Homologous Recombination by DSS1 via RPA Targeting and DNA Mimicry. *Mol. Cell* 59, 176–187. doi:10.1016/j.molcel.2015.05.032